DOI: 10.15740/HAS/AU/12.TECHSEAR(2)2017/305-327
Agriculture Update
Volume **12** | TECHSEAR-2 | 2017 | 305-327
Visit us : *www.researchjournal.co.in*
e ISSN-0976-6847

RESEARCH ARTICLE :

Molecular insights of floral malady prevailing in Indian soybean (*Glycine max* L.)

P. V. JADHAV, P.B. KALE, S. KAD, M.P. MOHARIL, M.S. DUDHARE, A.G. DESHMUKH, S.S. MANE, J.G. MANJAYA AND R.S. NANDANWAR

ARTICLE CHRONICLE :
Received :
10.07.2017;
Accepted :
23.07.2017

SUMMARY : Molecular alterations of symptomatic and asymptomatic soybean plants in response to floral bud distortion (FBD) are not completely understood yet. Oilgo-dT anchoredcDNA-RAPD and cDNA-SCoT markers were exploited to determine differentially expressed genes in response to FBD. Differential transcriptome analysis using cDNA based oligo-decamer profiling was executed to identify differentially expressed TDFs. The TDFs were further analyzed using computational tools *viz.*, BLAST homology, protein interactome, virtual karyotyping, sub-cellular localization, neighbor joining placement.The *in-silico* studies were used for annotation, structural, functional characterization and protein interaction of TDFs. Differential expression studies (cDNA-RAPD) produced transcript derived fragments (TDFs) in all tissues (leaf bud and node) of symptomatic and asymptomatic plant. Out of 197, only 26 TDFs were found differentially regulated. Amongst them 15 were found completely polymorphic and 11 showed differences in their amplicon intensity. Similarly cDNA-SCoTit revealed that total of 36 primers amplified 86 fragments between 200 bp to 1800 bp in length. Amongst them, nine differentially expressed fragments (DEFs) were re-amplified and sequenced. Sequence based studies of DEFs revealed their homology to five known genes; which were functionally involved in DNA repair mechanism, apoptopic, autophagy and pathogen responsive mechanisms in soybean. However, four fragments encode un-clarified proteins with unknown functions. The possible role of unknown genes, along with protein-protein interactome, physical karyotyping and cellular localization were predicted using *in silico* analysis. This study suggests that oligo-dT anchoredcDNA-RAPD, cDNA-SCoT differential display approach is a useful tool to serve as initial step for understanding alterations involved in upcoming malady and provide valuable information for further studies.

KEY WORDS :
Soybean, Floral bud distortion, cDNA-RAPD, cDNA-SCoT, Differentially expressed fragment

Author for correspondence :

P. V. JADHAV
Department of Agricultural Botany, Biotechnology Centre, Post Graduate Institute, Dr. Panjabrao Deshmukh Krishi Vidyapeeth, AKOLA (M.S.) INDIA
Email: jpraveen26 @yahoo.co.in

See end of the article for authors' affiliations

How to cite this article : Jadhav, P. V., Kale, P.B., Kad, S., Moharil, M. P., Dudhare, M. S., Deshmukh, A.G., Mane, S.S., Manjaya, J.G. and Nandanwar, R.S. (2017). Molecular insights of floral malady prevailing in Indian soybean (*Glycine max* L.). *Agric. Update*, **12**(TECHSEAR-2) : 305-327; **DOI: 10.15740/HAS/AU/12. TECHSEAR(2)2017/305-327.**

BACKGROUND AND OBJECTIVES

Flowering is a crucial step in the process of plant reproduction, the recognition of a compatible pollen grain, as it directly impacts fertility. Recent years, it is demonstrated that soybean has been facing a serious set-back due to occurrence of 'floral bud distortion' (FBD) consistently in central India. The average incidence of the floral bud

distortion was ranged from 8.0 to 14.6 per cent in different parts of central India and it was irrespective of genotype (Jadhav *et al.,* 2013). It is a condition where plant fails to produce pod and does not reach the maturity stage and continues remaining green at the end of season. The similar symptoms have been characterized as pod set failure syndrome (Rahimian *et al.,* 1995), bud blight (Arunkumar and Chendrayan, 2008), no-podding syndrome (Bhatia *et al.,* 1999) and bud proliferation syndrome (Lee *et al.,* 2011). Viral etiology was suspected through the reactions of selected plants in previous reports (Arunkumar and Chendrayan, 2008 and Golnaraghi *et al.,* 2004). Attempts to understand the disorder is limited because of inadequate knowledge of typical symptoms and causes. The symptoms produced did not completely resemble with any of the documented diseases, pests or disorders. Plants showing typical symptoms are found distributed inconsistently across locations every year (Jadhav *et al.,* 2013). The disorder is associated with deformities in flower and subsequent seed development due to inhibition of pollination and fertilization; leading to significant yield loss in symptomatic plants (Pracros *et al.,* 2006). Reports in this aspect are few but despite an upturn in evidence of losses, most researchers, extension workers and growers are still unfamiliar with the disorder (Jadhav *et al.,* 2013). However, many questions are remained in the field as the identification of cause(s) and cellular targets associated with the disorder.

Similarly, during successful plant development there are numerous molecular events regulated by interacting with surrounding environment. A complex interaction in plant metabolic networks is an outcome of fine regulation of plant genes singly or synergistically. The result of developmental alteration like FBD is projected as a partial or complete yield loss. Stepping along with the improving developments in the field of molecular biology, many new promising alternative techniques have emerged. To determine molecular alterations, different approaches have been used in different crops which are needed to explore towards understanding nature of FBD prevailing in soybean.Currently, many approaches have been developed for the analysis of differential gene expression at the mRNA level in various plants. Some of the most notable methods include mRNA differential display (DD), representational difference analysis (RDA), suppression subtractive hybridization (SSH), serial analysis of gene expression (SAGE), cDNA-AFLP, cDNA-SRAP, cDNA microarray, etc. (Velculescu *et al.,* 1995; Bowler, 2004; Blackshaw *et al.,* 2007; Huang *et al.,* 2007a and b; Liang *et al.,* 2007; Nettuwakul *et al.,* 2007; Polesani *et al.,* 2008; Wee *et al.,* 2008; Hillmann *et al.,* 2009; Zamharir *et al.,* 2011 and Xu *et al.,* 2011). However, regardless of the abundant quantity of available methods, there happens to be both advantageous and disadvantageous traits for each system. To study the differential expression of genes and to see the epigenetic changes involved in the plants various molecular markers were used. These include cDNA-RAPD (Nimbalkar *et al.,* 2006), Start Codon Targeted (SCoT) markerbased on the short conserved region flanking the ATG start codon in plant genes (Collard and Mackill, 2009).

Objectives :

– Differential expression studies in floral bud distorted and asymptomatic soybean plants at molecular level using cDNA RAPD, cDNA SCoT

– Insilco studies using bioinformatic tools.

RESOURCES AND METHODS

Plant material and sampling stage :

The experimental material comprised of symptomatic and asymptomatic soybean plants collected from the experimental field of Department of Agricultural Botany, Dr. Panjabrao Deshmukh Krishi Vidyapeeth, Akola, Maharashtra State, India (20.42°N) during 2014-15. Collection of plant samples of a popular Indian soybean genotype, JS-335 was made at R6 growth stage (pod containing green seed that fills the pod cavity at one of the four uppermost nodes with completely unrolled leaf). However, the plant tissues (node, stem, root) were collected in liquid nitrogen and brought to laboratory. All samples were maintained at -80°C and used for molecular profiling.

RNA isolation and cDNA synthesis :

Symptomatic and asymptomatic soybean tissues were collected from three replicated samples. Total RNA was extracted from the frozen nodal tissues (R6 stage) using the Pure Link ® RNA Mini Kit (Ambion Life Technologies Corporation, Carlsbad, USA) and stored in 100 µl of elution buffer (10 mMTris-Cl, pH 5.8). Reverse transcription of transcripts and second-strand synthesis from 500 ng of total RNA was carried out using Moloney murine leukemia virus (M-MLV) reverse

transcriptase (Invitrogen, Cat. No. 28025-013) and quantified by measuring OD at 260/280 and 260/230nm.

cDNA-RAPD profiling :

First strand cDNA generated were equalized and used as a template for further amplification studies. Amplification was carried out by using a Thermal cycler (Eppendorf Mastercycler gradient No: 5331 and Eppendorf Mastercycler personal No: 5332) using number of RAPD primers using 2X Master Mix (Fermentas), *Taq*DNA polymerase (Fermentas); genomic DNA as per the standard protocol.PCR amplification in all tissues (leaves, pod and node) from both symptomatic and asymptomatic (control) plant were achieved using RAPD primers. As much as 20µl of PCR reaction mixture contained 10ng of cDNA, 2.0 µl 10X PCR buffer, 2.0mM $MgCl_2$, 2mM dNTPs, 10ng of primer and 1 unit of *Taq* Polymerase (Ferments) was used. Amplifications were performed by a cycles of: 5.00 min at 94°C followed by 39 cycles each of 1 min at 94°C, 1 min at 36°C and 2 min at 72°C and final extension of 15 min at 72°C. The PCR products were resolved on 1.8 per cent agarose using TBE buffer.

cDNA-SCoT profiling :

The PCR amplification of quantified second-strand cDNA from both symptomatic and asymptomatic plant samples (at R6 growth stage) were achieved using oligo-dT anchored start codon targeted markers (SCoT). As much as 20 µl of PCR reaction mixture containing 10ng of cDNA, 2 µl of 10X reaction buffer, 1 µl of RiboLock, 2 µl of 10 mM dNTP mix, 1 µl of 20 pmol primer and 2 units of M-MLV reverse transcriptase. Amplification were performed for 41 cycles: an initial denaturation at 94°C for 5 min followed by 40 cycles, each of 1 min at 94°C, 1 min at 48-55°C and 1 min at 72°C followed by final extension for 7 min at 72°C. The PCR products were resolved on 2.5 per cent agarose in 1X TBE-buffer. All the reactions were repeated at least three times and consistently reproducible amplicons were scored.

Scoring amplicons and tissue wise clustering :

Further the recording for differential analysis on the basis of number of amplicons (present/absent) as well as differences in amplicon intensities was carried out to understand differential expression pattern in symptomatic and asymptomatic soybean tissues in all of the three molecular markers.

Amplicons derived binary data in the form of 0 (for absence) and 1 (for presence) is further used to generate dendrogram and to study methylation pattern associated with floral melody using MSAP marker. This is used to create groups of the individuals according to Jaccard's coefficient used to compare between set of variables. Here this type of grouping was used to distribute tissues under investigation.

Elution and sequencing of differentially expressed fragments (DEFs) :

The individual DEFs were eluted from agarose gel with sharp surgical blade without contaminating other fragments using QIAEX II gel extraction kit (Quiagen Inc., Valencia, CA). Aliquot of 1 µl was used for re-amplifying individual DEFs using same corresponding primer and analyzed electrophoretically. The PCR-amplicons were identified by nested PCR with corresponding primers and submitted to the Eurofins Genomics India Pvt Ltd., Bangalore for custom sequencing.

Sequence characterization of DEFs :

The five sequences of DEFs were analyzed for their homology using Chromas Lite 2.01 software. The sequence similarity analysis was done using BLAST homology against the publically available nonredundant genes/ESTs/transcripts/proteins using BLAST algorithms (*http://www.ncbi.nlm.nih.gov /Blast.cgi*) (Altschul *et al.*,1997). Sequences were subjected for their possible computational characterization and functional prediction and properly annotated one sequence was deposited to the NCBI.

***In-silico* protein interactome and sub-cellular localization of DEFs :**

Deciphering the interactive links between proteins is needed to understand its role in multifaceted metabolic pathways. Reconstruction of complex interacting pathways integrating predicted interaction networks with available experimental data is becoming one of the most demanding requirements in the post-genomic era. This method can address the position of physically interacting proteins in pairs and identifies the most likely motifs involved in the interactions. Therefore, sequences annotated after BLAST homology were subjected to this

study by protein-protein interaction tool-STRING as described by Franceschini *et al.* (2013). This is useful to discriminate between true and false interactions in a significant number of cases, so as to generate information of target proteins. Here, virtual reconstruction of complete interacting networks can be determined to understand the possible role of protein (s) associated with FBD.

Sub-cellular protein localization :

Sub-cellular localization of uncharacterized sequences was carried out using WoLF-PSORT a web based tool. Further, this was used to obtain their possible role in the plant metabolisms. Information generated for uncharacterized proteins will help to understand their sub-cellular locations and possible role in molecular alterations associated with the FBD.

***In-silico* physical karyotyping of DEFs :**

The available soybean databases (Phytozome, SoyBase and NCBI) were used to anchor identified DEF sequences generated on virtual chromosomes of *G. max* to identify their distribution, relative position and abundance. The exact locations of DEFs were determined using MegaBLAST tool showing at least 80 per cent identity according to Soares-Cavalcanti *et al.* (2012).

OBSERVATIONS AND ANALYSIS

The results obtained from the present study as well as discussions have been summarized under following heads and Tables 1 to 10 and Fig. 1 to 3.

Signs and symptoms of floral bud distortion :

The floral bud distortion is a condition where soybean plant fails to produce pods and doesn't show senescence at the end of the season. The noticeable morphological symptoms recorded were the plants with prolonged vegetative phase (stay green) even after R8 stage without pod development. The symptomatic plants were found randomly distributed in the field and failed to produce pods leading to complete yield loss. The symptoms produced did not completely resemble with any of the documented diseases or pest of soybean. The disorder mainly affects flower development and leads to either no or deformed buds and pods. The disorder depicts proliferation of floral bud followed by extended vegetative phase in severely affected plants. At later growth stages plant show failure in pod set, hence only infrequently developed pods may appear. The floral bud distortion exhibited decreasing apical dominance and distorted floral organs *i.e.* petals, stamens, carpel and ovary, leading to failure of pod development on affected plants (Fig. 1).

Molecular profiling :

cDNA RAPD profiling for TDFs :

Amplicon / TDFs generated in the study :

Out of total 197 amplicons produced in the experiment described earlier, 173 amplicons were monomorphic and 26 were found differentially regulated. These were recorded either completely polymorphic or with varying intensities. Out of 26 differentially expressed Transcript Derived Fragments (TDFs), 15 were found completely polymorphic and 11 showed differences in

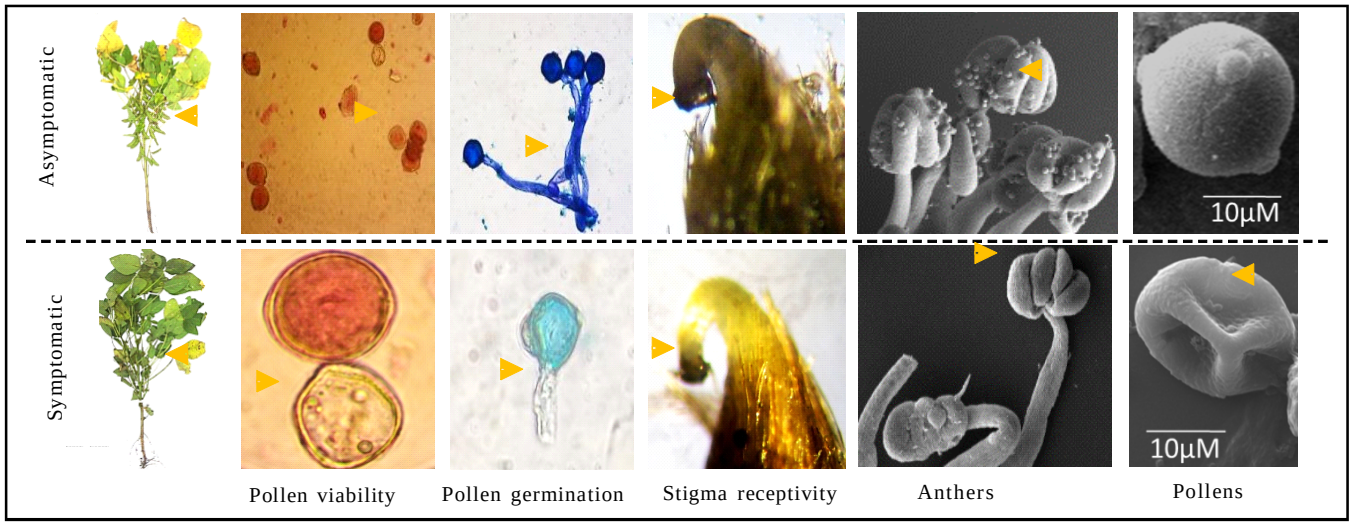


Fig. 1 : Cytological behaviour of floral reproductive organs of symptomatic and asmptomatic soybean plant

their amplicon intensities (Table 1).

Computational analysis using bioinformatics :

In the present investigation, various *in silico* analysis tools were exploited to characterize the TDFs derived in cDNA RAPD experiment. Since; TDFs are representatives of differentially expressed genes in FBD symptomatic over asymptomatic plants. Hence, for TDF, their detailed annotation, relatedness with database entries, functional characterization, protein interactions and their sub-cellular locations are necessary. This information will help to trace out their possible involvement in floral bud distortion. Freely available online computational tools were used in this experiment.

Sequencing and homology analysis of TDFs :

The individual differentiating TDFs were and re-amplified using 1µl of eluted amount and respective primer. Amongst differentially expressed TDFs those found reproducible and recorded with an expected range were re-amplified. These were processed in custom sequencing services.

Sequencing results of only nine TDFs were found decipherable and edited using Chromas Lite 2.01 software in lab. The final sequences generated were copied as plain text and evaluated individually by online BLAST homology search tool (*http://www.ncbi.nlm.nih.gov*) (Altschul *et al.,* 1997).

The homology search of TDF sequences against the public database used to identify the cellular functions on the basis of maximum similarity with existing well characterized genes. Thus, we can assign putative functions to isolated genes. In preset investigation, majority of the TDFs were found involved in plant developmental processes e.g. plant cell rescue, defense, signaling etc. Amongst nine TDFs generated in this investigation, five were found to be up-regulated and four were down-regulated. These information and sequences were further studied with *in silico* interactome tool (Table 2)

cDNA-SCoT profiling for differentially expressed transcript derived fragments :

Analysis of gene expression is a central aim in most studies of molecular and cellular biology. It forms the basis for unraveling the control of plant growth and development. Also it allows the identification of specific regulatory key points of metabolism (Gupta *et al.,* 2013). SCoT polymorphism is a gene-targeted marker that can generate information interrelated with biological traits compared with random DNA markers (Collard and Mackill, 2009 and Xiong *et al.,* 2009). The cDNA-SCoT technique has been explored in differential expression studies such as, Wu *et al.* (2010) analyzed the differential expression of gibberellin-induced genes for stalk elongation of sugarcane using cDNA-SCoT. Considering the potential of SCoT markers, present investigation was intended to study differential expression in reference to floral bud distortion in soybean. The quantified cDNA of nodal tissue of both symptomatic and asymptomatic plants (R6 stage) was profiled using oligi-dT anchored SCoT markers. Altogether 36 primers produced 92 scorable amplicons ranging from 200 to 1800 bp. Out of them, 74 were monomorphic and 18 were differentially expressed. The maximum of 24 amplicons were produced by SCoT-18 primer followed by SCoT-19 (20), SCoT-17 (11), SCoT-13 (8), SCoT-14 (7), SCoT-12 (5), SCoT-15 and 16 (4 each) and SCoT-1 (3) SCoT-21 and Study of SCoT-32. Amongst differentially expressed amplicons, 5 were reproducible including 4 up and 1 down regulated in symptomatic plant. These five amplicons were eluted and sequenced. The differentially expressed fragments were characterized with the help of *in-silico* tools.

Table 1: Classification of differentially expressed TDFs

Class	Origin of TDFs	No. of TDFs	Remark/ significance	
A	Amplicongenerated only in symptomatic tissue (FBD)	10	Represents genes responsible for FBD	OPD1, OPF17, OPF16, OPF20 OPD12, OPF14, OPF12,OPA05,OPA16,OPA08
B	Amplicon showing variation in intensity	11	Represents genes up-regulated in FBD	OPD2, OPD12/12, OPD18, OPF14/14, OPF12/12, OPF15,OPA04
C	Amplicongenerated only in asymptomatic tissue (Healthy)	05	Expressed and up-regulated in healthy plant	OPF16, OPF17, OPF20, OPA13, OPA17
	Total differentially expressed TDFs	26		

Table 2 : BLASTX studies of the sequences homology using Leguminosae (taxid: 3803) organism and Ref Seq_RNA database

Sr. No.	Name assigned	Regulation pattern in FBD	Sequence	Length (bp)	Homology	Organism	*E* value	Percent identity	Function	Reference
1.	OPA-03-SyP 475 partial	Up	TTCGACACAGGAACAAGTGGGAAATTGTCTTCACTGATGATGAAGGGGATATGATGCTTGTTCGACACAGGAACAAGTGGGAAATTGTCTTCACTGATGAAAGAATCTTCATTTGTTCAAGCCAGGATGTGCACAAATTGAGCTCAGGGAGCAAACTACCTATCTCTTCTATGGGAGAGATTGTAATAAGCTTAAACACAACCGAGACCTGAAACTTTTACATTCTTGTTCTTACTATTTTGTTGGTTTTTTTTCTTTTATTTTTCTGTTGCTTTTATTTTGCTTATTCAAGGTAAGAATGTTAGTTTGTTTTTGGCTAGGCAAG	325	PREDICTED: Auxin response factor 9-like (LOC100799088), mRNA	*G. max*	2e-4	65/78 (83%)	Auxin response factors (ARFs) are transcription factors that bind to TGTCTC auxin response elements	Gray *et al.*, 1999 Hagen and Guilfoyle, 2002 Ellis *et al.*, 2005
2.	OPA-16-AllAs Partial	Down	CTAATCAAAGTCATGAGCAAAAACTCAATGGGAAAGGAGGGGTTAATGAAATCACTCCAAATTATTTGTTTATTCTCTAAATTAAATTTTAAAAAATCTTGGTCCAAAATTATCAGTAATTGTTTTTAAATGATGGAGGTGATTATCTTCCGTTACTAGCTAGAGTTTACTTTTCTTTCCTTTTCTCTTTTGATAAATTTTTTATATATAAAAGTTATGTCATTGTTTTATTTATCATTTTCAGATCTCTGTTTCTGGTCAAATTCAGTTTGAGCCTCGAAAAACCTATGTCTACATTTATCCA	304	PREDICTED: Uncharacterized LOC100815325 (LOC100815325), transcript variant X1, mRNA	*G. max*	3e-32	62/80 (78%)	Unknown function	-
3.	OPD12_700B- partial	Down	ACGAGAGCAGGAACTAGGGGAAGGAAGAAATTGCAGAATGAGCCACCATTGGATAGCAGTACAGTTATCATAATTGGAACATATTGAAAACGTCGAGGAGAAGAGCTTAAGGAGGAAATAAACCTGTTGACGTGGGAGAAGGTAAGAAAAATGCTAATTTTGACGAAAGATGCAGTGGGTCGAGCCCTCAAGAGGTTTGTGATCGAGATGAAAACCAAGATGGTGTATCATTTCAGAAGGTTGCGAAGAGGTTGCTGATGGGCGAGCTTCACATGATGAGGACTTCCTTTGTAAGCTGAGAAGGCGCCGGATTTGAAGAAGATGACACTTGGAGATTGGTTTGATTATTTGGAGACTCATTTGCCTAGACAAATTATTGATGCAACCGAAGAGGTAATTTCTGGCATGAAAATTAAATCTAAACAAGTACAGGAGTATGTTGTTCAGCAGAAGATTGAGAATTGCCAAGGAGATTAG	477	PREDICTED: FHA domain-containing protein At4g14490-like (LOC102667282), mRNA	*G. max*	7e-15	30/52 (58%)	Involved in intracellular signal transduction, cell cycle control, transcription, DNA repair and protein degradation	Morris *et al.*, 2006
4.	OPF-19SyC-300_Partial	Up	ATTTTGAACCACTCCCCGATATAATCTAACATTACATTTGATATTTACTTATATATATATTTAAAAATATATTATACATTAAATTAGTTCTATTAATCGGTATCTTCTAGTTCTAGTCTGCGACATGGATAATAACTAGGATACTCGAGGGAGGCTGCGGTCAATTTTACCGTTGAAAGAATGAAAGTTATGCTCTCTTTTTTTGTATATATGAAAGTTATGTCTTTTGCTTCCATGAAAGTTCACATATGACATGCCAATGATCAGAATCAGATGCTCCCTGATCGGTCTAGAGG	296	PREDICTED: Probable isoaspartyl peptidase/L-asparaginase 2-like (LOC101492649), mRNA	*C. arietinum*	1.4	10/25 (40%)	Nitrogen supply to sink tissue in legume developing seeds	Lurin*et al.*, 2004

Table 2 : Contd.....

Table 2 : Contd.....

5.	OPF-19-400-B_Partial	Up	CCGAAAACCACCCCGATAAAAAATGCTCCACTAAA ATTGATATTTACTTATATCTATATTTACCCCATCTTA TACAGTTAACTTGTTTTTATACTCGGCTCTTCAAAT TCTAGCCTGCGATATGGATAATAAATAGGATACTC GAGGGAGGCTGCGGTCAATTTTACCGTTGAAAGAA TGAAAGTTATGCTCTCTTTTTTTGTATATATGAAAG TTATGTCTTTTGCTTCCATGAAAGTTCACATATGAC ATGCCAATGATCAGAATCAGATGCTCCCTGATCGG TCTAGAGGC	294	Hypothetical protein (MTR_4g090180) mRNA, complete cds	*M. truncatula*	0.19	14/30 (47%)	Hypothetical protein	-
6.	OPD15-400-A_Partial	Down	TCAGATCTTATTTTTTGAAAATTGAAACGAGCGAA AAGGTAAATCGCGTTGTTTGGATGTTCAACGGGCA ATCGGTTCTATGGGGTTTTTTAGTCAGAAAGGCTAG ACAAGTCGGGTGAGTGCTGTGGTGCCATAGATAAA GATTGATCCCCATAAAAAAGAAATCAGGCTCTTCT GCGAACCGTAAGAAGATTAATTATTTAAGCCACCT CGTCCGTGTCTAGATGTTTGAAGACCGCATTAAACT CCTGAGTTGTACAAGTACTAATGGCCGCGACGGTT ACACGTCCACCCCCCCTTCATACTGTTCTACCATG	317	Hypothetical protein (PHAVU_005G068000g) mRNA, complete cds	*P. vulgaris*	1.6	19/19 (100%)	Hypothetical protein	-
7.	OPF12-400-A_S_Partial	Up	CCCCGGATACCGGATAATTGTGCACCGTGGGGAAA AGGAAACGGCTTGCTTGGGTGTTGAAGGGGGAATT GGTGAGTTTAAAGAGTGTAGCAAGGTAGGGGAGA GGGTTGGGGGGGGGATGTTGGGGAGCCGATGGCAA GGGTTGGACCTTATGCAAACGAATTCCAGACCATC CTGCGCACTCTGCGAAGATTATCGACATTATGCTAC CACGCGGGCTTCTGCAATTTTCACACGGCGTAACC CCCCGAGGTACCCCAAGACTAAAATGCCTGTCGAT TTCACTTCCTCACCCCCCTCCCACCCTTTGAACCCA	315	PREDICTED: Pentatricopeptide repeat-containing protein At3g22690-like (LOC100798950), mRNA	*G. max*	0.033	16/36 (44%)	restores fertility to cytoplasmic male-sterile plants	Cho *et al.*, 2007
8.	OPD12-A_250_ Partial	Down	TCGGCAAATCCCTTAATAGAGGAGTGCAAGTACAT GTGATCGGCTTATATTGGTGTTTTACAGAGCATCTG TTCTACGCGGCAAAAAGCACAGTGACTCAGAATAA GGTAAGCGAACCCCAAACCCTTCTGATTTACCCAC CCTACAAATCTCATGCGCACAACCTC	167	Hypothetical protein (PHAVU_002G071200g) mRNA, complete cds	*P. vulgaris*	0.78	27/31 (87%)	Hypothetical protein	-
9.	OPD13 A_Partial	Down	ACGGGCGGACGCTTAACAAAGGAGCGCTCGGGAC ACTTAAACGGCTTGTCTGCGGTGTTCCAGACTCACT TTGGTGACAGCTAGTTTTTTTGCACTTAGGCCAAAC GGATCATGTAAACAGGGTCATCCCAGACACCCAAG GTAGACACCCTATAAAAAAAAAATCGAGGACCCCCC TTCCGCCTCGCAATGAATTTTCTCGGTTTTGCAGAC CTCGCGGTCCTGAAGCGGTTTTAACCCCGCGTTTGA AGCACCTTTGTAGTTACTAATCATTAGGCCCCGACG ATTCGACGCCCTCCCCACCTTCCTACT	311	PREDICTED: Uncharacterized LOC102667978 (LOC102667978), mRNA	*G. max*	1.6	21/22 (95%)	Uncharacterized protein	-

Sequence homology and characterization of DEFs using *in-silico* tools :

The sequence comparison of 5 differentially expressed fragments (DEF) against the database towards identifying cellular process (altered due to floral bud distortion) revealed homology to genes with known functions. These DEFs were found involved in plant developmental processes particularly, in programmed cell death, DNA damage repair mechanism, apoptopic and autophagy revealed programs, response to pathogen infection and post translational modifications. These may represent potential transcripts involved in floral bud distortion and need to be further studied.

Comparative analysis was undertaken across species for correlating protein-protein interaction to provide an insight in to the regulation of floral bud distortion at a molecular level (Fig. 2A). Amongst 5, only one DEF showed possible role in floral development. The DEF generated from the marker SCoT-14 showed homology with the PARP gene, responsible for DNA damage repair and programmed cell death (PCD) in plants (Rogers, 2005). In case of floral bud distortion, multiple viruses are reported to be found in symptomatic plants (Jadhav *et al.*, 2013) and hence, plant may show stress response which leads to abnormal morphology.

Overall plant may receive mild stress and hence PARP gene expressed in, may play role of DNA damage repair in all plant parts including reproductive organs *i.e* pollen. As compared to overall plant parts, floral organs are more prone to any stress (Koti *et al.*, 2005). In case of pollen developmental stages, PARP gene may act as inducer of plant cell death. Therefore, the hypothesis correlates with the presence of expression of PARP gene in symptomatic plant. However, remaining 4 DEFs were un-characterized protein; therefore we tried to determine their sub-cellular location using *in-silico* tools.

Computational analysis of DEFs using *in-silico* tool:

DEFs are representatives of differentially expressed genes associated with FBD in symptomatic over asymptomatic plants. Hence, the DEFs were characterized for their annotation, relatedness with database, functional characterization, protein-protein interactions and their sub-cellular locations. This information will help to trace out their possible involvement in floral bud distortion. In this context, freely available online computational tools were used and their details are described in each of the subheading below.

Sequencing results of only two DEFs were found decipherable and edited using ChromasLite 2.01 software. The final sequences generated were copied as plain text and evaluated individually using online BLAST homology search tool (*http://www.ncbi.nlm.nih.gov*) and results generated were depicted in (Table 3).

Sub-cellular localization and functional characterization of the un-annoted proteins :

The uncharacterized proteins were screened to understand their cellular location and possible role in the plant metabolisms. Numerous experiments were used to determine protein localization, bilipid membranes divide eukaryotic cells into various organelles containing characteristic proteins and performing specialized

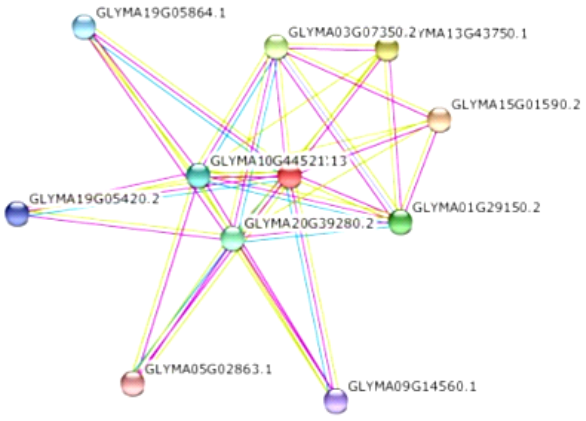


Fig. 2A : Networking of PARP with other plant proteins

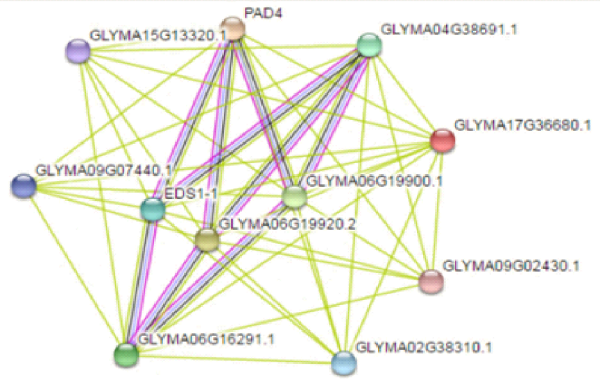


Fig. 2B : Thaumatin-like protein with other plant proteins

Table 3: BLASTX studies of the sequences homology using *Glycine max* (taxid: 3847) organism and non- redundant database

Sr. No.	Name assigned	Regulation pattern in FBD	Sequence	Length (bp)	Homology	Organism	*E* value	Percent identity	Function
1.	SCoT-21: 350bp	Down	CTAATCAAAGTCATGAGCAAAAACTCAATGGGAAAGG AGGGGTTAATGAAATCACTCCAAATTATTTGTTTATTCT CTAAATTAAATTTTTAAAAATCTTTGGTCCAAAATTATC AGTAATTGTTTTTTAAAATGATGGAGGTGATTATCTTCCGT TACTAGCTAGAGTTTTACTTTTCTTTCCTTTCTCTTTTGA TAAATTTTTTATATATAAAAGTTATGTCATTGTTTTATT TATCATTTTCAGATCTCTGTTTCTGGTCAAATTCAGTTT GAGCCTCGAAAAAGCCTATGTCTACATTTATCCA	304	PREDICTED: Uncharacterized LOC100815325 (LOC100815325), transcript variant X1, mRNA	*G. max*	3e-32	62/80 (78%)	Unknown function
2.	SCoT-31: 500bp	Up	CCGAAAACCACCCCGATAAAAAATGCTCCACTAAAATT GATATTTACTTATATCTATATTTACCCCATCTTATACAG TTAACTTGTTTTTATACTCGGCTCTTCAAATTCTAGCCT GCGATATGGATAATAAATAGGATACTCGAGGGAGGCT GCGGTCAATTTTACCGTTGAAAGAATGAAAGTTATGCT CTCTTTTTTGTATATATGAAAGTTATGTCTTTTGCTTCC ATGAAAGTTCACATATGACATGGCAATGATCAGAATCA GATGCTCCCTGATCGGTCTAGAGGC	294	Hypothetical protein (MTR_4g090180) mRNA, complete cds	*M. truncatula*	0.19	14/30 (47%)	Hypothetical protein

Table 4: List of un-annotated proteins and their sub-cellular localization

Sr. No.	DEF	Regulation	Homology	Organism	Function	Localization
1.	SCoT-14 As-1	Up- regulated in asymptomatic plant	*Arabidopsis thaliana* protein pleiotropic regulatory locus 1 mRNA, complete cds	*Glycine max*	Essential for plant innate immunity.	Nucleus
2.	SCoT 13-S-1	Up- regulated in symptomatic plant	Auxin-binding protein 1	ABP- Maize	Receptor for the plant hormone Auxin	Endoplasmic reticulum
3.	SCoT-14 S-2	Up- regulated in symptomatic plant	*Arabidopsis thaliana* chorismate mutase 1 mRNA, complete cds	*Arabidopsis thaliana*	Physiological role in developmental and environmental signals of plant	Nucleus
4.	SCoT-15 S-1	Down regulated in symptomatic plant	*Arabidopsis thaliana* RP3B	RP3B_ARATH	Catalyzes the transcription	Cytosol

functions. Thus, subcellular localization information provides important clue of protein's function (Horton *et al.*, 2007). In this exercise, the information about sub-cellular localization (WoLF PSORT) along with protein description was used in functional classification of unknown proteins found. Four uncharacterized proteins representing DEF's showed varied locations like nucleus, cytosol with their different functions (Table 4).

DEF involved in programmed cell death:

Poly (ADP-ribose) polymerase 3-like (*Glycine max*):

The circadian clock is a timing mechanism by which plants co-ordinate the temporal phases of physiological processes to daily changes in their environment resulting from earth's rotation. Various studies have established the role of circadian oscillator in photoperiodic control during flowering (Samach and Coupland, 2000). Extreme conditions such as floral bud distortion have a strong impact on plant growth, development and effects are attributed by interconnected signaling pathways like hormones and metabolism, both of which regulate growth even under normal conditions.

Here we demonstrate that the DEF generated by SCoT 14-As-1 was homologous to PARP gene, and has been implicated as one of the enzymes in the apoptotic pathways persuaded by DNA damaging agents or oxidative stress. Apoptosis in plants plays important roles in phases of development (embryonic, seedling and maturity. Also apoptosis often observed during the plant response to pathogen attack, where it is characterized as 'hypersensitive

response' (HR) (Jones and Dangl, 1996), apoptic-like plant cell death in floral organs is in the death of pollen-tube during self- incompatible pollination interactions. Similarly in Paper, Thomas and Franklin-Tong (2004) showed that self-incompatibility (SI) stimulated due to increase in cytosolic (Ca^{2+}), which in turn activates release of cytochrome-c into the cytosol and induced caspase- 3-like activity. It has been observed that, caspase-3-like activity cleaved poly(ADP-ribose) polymerase (PARP), a classic substrate for caspase-3 enzymes, and was correspondingly inhibited by the peptide Ac-DEVD-CHO, which blocked DNA fragmentation and pollen-tube growth.

Similarly, PARP are amongst the interacting proteins which having functions like Poly [ADP-ribose] polymerase 3,BTB/POZ and TAZ domain-containing protein 3, Transcription factor bHLH140-like (762 aa), ATP-dependent DNA helicase 2 subunit KU70 and which are involved in various plant developmental stages. One of the interacting proteins of PARP is bHLH transcription factor which plays an important role in pollen development. These proteins belong to family of transcription factors that bind to DNA targets as dimer (Toledo-Ortiz *et al.*,2003 and Heim *et al.*, 2003).Such factors involved in symptomatic plant tissue hence, bHLH down regulates this protein activity. As discussed earlier, PARP activity is involved in the regulation of the cell cycle for development by influencing a particular subset of genes, as well as energy homeostasis, primary and secondary metabolism leading to enhanced growth. This data provides new insights in to plant growth regulation, shows that PARP could be one of the prominent players of floral bud distortion and opens interesting new starting point to understand the distortion at molecular level.

Protein kinase PVPK-1-like (*Glycine max*):

In this study, a DEF named "SCoT 13-S-2" (Accession number KT619134) was found to be up-regulated in symptomatic nodal tissue. The *in silico* studies revealed most relevant hit with *Arabidopsis thaliana* protein with pleiotropic regulatory locus 1 mRNA which encodes a nuclear WD40 protein imported into the nucleus. It is essential for plant innate immunity.

It is involved in variety of functions ranging from signal transduction and transcription regulation to cell cycle control and apoptosis. Repeated WD40 motifs act as a site for protein-protein interaction, and proteins containing WD40 repeats are known to function as platforms for the assembly of protein complexes or mediators of transient interplay among other proteins (Zhang and Zhang, 2015). It plays important role in flowering autonomous pathway which positively regulates flowering by promoting transcriptional repression of flowering repressor Flowering Locus-C gene (FLC). The process with specific outcome is progression of the flower over time from its formation to the mature structure. The flower is the reproductive structure in a plant, and its development begins with the transition of the vegetative or inflorescence meristem into a floral meristem.

DEF involved in responses to pathogen infection : *Protein YLS9-like* [*Glycine max*]:

In this study, a DEF derived from "SCoT 13-S-1" was found to be up-regulated in symptomatic nodal tissue. *In silico* studies revealed UniProt link for the most relevant hit showed homology with Auxin-binding protein 1 found in nucleus**.** A series of molecular signals produced by the binding of the plant hormone auxin to a receptor, and ending with modulation of a downstream cellular process of transcription. Hesse *et al.* (1989) reported that protein YLS9 acts as receptor for the plant hormone auxin. An YLS9 cDNA (866 bp) contained an open-reading frame encoding a polypeptide of 227 amino acid residues, which was similar to hin1 of *Nicotianatabacum*, a hyper - responding gene on pathogen infection (Gopalan *et al.*, 1996).

PREDICTED: *Glycine max* thaumatin-like protein 1 (LOC106798928), transcript variant X2, mRNA:

DEF showed homology with PREDICTED: Glycine max thaumatin-like protein 1 (LOC106798928), transcript variant X2, mRNA (Fig. 2B).

Predicted function and role of gene: Thaumatin-like proteins:

Thaumatin-like proteins (TLPs) are the products of a large, highly complex gene family involved in host defense and a wide range of developmental processes in fungi, plants, and animals. Despite their dramatic diversification in organisms, TLPs appear to have originated in early eukaryotes and share a well-defined TLP domain (Liu *et al.*, 2007). Thaumatin-like proteins (TLPs) are polypeptides of about 200 amino acid residues

that share sequence similarity with thaumatin (Velazhahan *et al.*, 1999). TLPs are classified as the pathogenesis-related (PR) protein family 5 (PR5), 1 of 17 families of defence-related PR proteins (Christensen *et al.*, 2002 and Van Loon *et al.*, 2006) pathogenesis-related (PR) proteins which represent major quantitative changes in soluble protein during the defense response. The PRs have typical physicochemical properties that enable them to resist to acidic pH and proteolytic cleavage and thus survive in the harsh environments where they occur vacuolar compartment or cell wall or intercellular spaces (Stintzi *et al.*, 1993).

TLP localization in polar cells and tissues is complex. TLP1 is expressed predominantly in tissues with a prominent vascular system such as midveins, petioles and stems, whereas the second TLP is primarily expressed in starch-storing plastids found in young leaves and the shoot apex. (Dafoe *et al.*, 2010).

String protein Interaction studies found that the target protein interacts with the regulatory proteins like regulatory protein NPR1-like GLYMA09G07440.1 NONEXPRESSOR OF PATHOGENESIS-RELATED PROTEINS1 (NPR1; also known as NIM1) (Table 5). Is a master regulator of systemic acquired resistance (SAR). SAR is induced by salicylic acid (SA), leading to the expression of PATHOGENESIS-RELATED (PR) genes.

Table 5: String protein interaction of PREDICTED: Glycine max thaumatin-like protein 1 (LOC106798928), transcript variant X2, mRNA

Predicted functional partners	
GLYMA06G19920.2	Uncharacterized protein (617 aa)
GLYMA06G19900.1	Uncharacterized protein (615 aa)
GLYMA06G16291.1	uncharacterized LOC100815981 (633 aa)
GLYMA04G38691.1	Uncharacterized protein (633 aa)
EDS1-1	Uncharacterized protein (612 aa)
GLYMA02G38310.1	Uncharacterized protein (432 aa)
GLYMA09G07440.1	regulatory protein NPR1-like (540 aa)
GLYMA15G13320.1	NPR1-2 protein; Uncharacterized protein (590 aa)
GLYMA09G02430.1	NPR1-1 protein; Uncharacterized protein (590 aa)

The NPRI nonexpressor of pathogenesis-related genes is a key gene involved in regulation of plant disease resistance. It plays a pivotal role not only in systemic acquired resistance (SAR) and induced systemic resistance (ISR), but also in basic resistance and resistance (R) gene-dependent resistance.

UTR characterization for the studies in the regulation of the gene by Ref Seq_RNA database of the two DEF which are non-redundant:

Sequencing results of two DEFs were found

Table 6: UTR characterization is necessary for the studies in the regulation of the gene by RefSeq_RNA database

Motif type	Motif name	Position	Length	Sequence
Human splicing sites	Acceptor	112 ~ 113	2	Ag
exon splicing enhancer (ESE)	tra2 beta1,	172 ~ 180	9	Gaaagaatg
	cftr, exon 12	136 ~ 142	7	Ggatact
polyadenylation sites (PAS)	Polyadenylation sites	138 ~ 169	32	Atactcgagggaggctgcggtcaattttaccg
Transcriptional regulatory motif	p53	247 ~ 256	10	Tgacatgcca
Transcriptional regulatory motif	GATA-1	12 ~ 21	10	Cccgataaaa
Transcriptional regulatory motif	Pax-4	3 ~ 14	12	Gaaaaccacccc
Transcriptional regulatory motif	LXRalpha:RXRalpha	151 ~ 165	15	Gctgcggtcaatttt
Transcriptional regulatory motif	TBP	88 ~ 94	7	Tttatac
Transcriptional regulatory motif	Kid3	8 ~ 12	5	Ccacc
Transcriptional regulatory motif	ZNF333	126 ~ 130	5	Ataat
Transcriptional regulatory motif	HOXA13	16 ~ 21	6	Ataaaa
Transcriptional regulatory motif	Cdx-1	122 ~ 137	16	Atggataataaatagg
Transcriptional regulatory motif	HNF-1alpha	72 ~ 88	17	Atacagttaacttgttt
Transcriptional regulatory motif	HNF-1beta	73 ~ 89	17	Tacagttaacttgtttt
Transcriptional regulatory motif	AP-2_gamma	111 ~ 117	7	Tagcctg
Transcriptional regulatory motif	MEF-2D	130 ~ 136	7	Taaatag
Transcriptional regulatory motif	CDX-2	88 ~ 93	6	Tttata

decipherable and edited using Chromas Lite 2.01 software. The results generated were depicted in (Table 6). The result showed UTR characterization which was studied further using RefSeq database.

Untranslated regions (UTR) play important roles in the post transcriptional regulation of mRNA processing. They function in various post-transcriptional events, such as mRNA turnover, polyadenylation, localization and translational initiation.Affymetrix single probes are reannotated and assigned to the coding region (CDS) or the 3'UTRof the transcript, according to NCBI RefSeqdatabase. For each RefSeq two distinct custom probesets, the first one including probes covering specifically the CDS and a second one including probes covering specifically the 3'UTR. UTRs contain motifs that are essential to regulate post-transcriptional processes (e.g. tra2 beta1, cftr, exon 12, ZNF333, HOXA13, GATA-1 etc). At the end of every mRNA there is a signal indicating that the end of the mRNA is reached called the polyadenylation (poly (A)) site. In many genes, two or more poly(A) sites are found in the 3'UTR, so that different isoforms with different 3'UTR length can be expressed.

Transcriptional motifs found in the PREDICTED: Uncharacterized (LOC100815325), transcript variant X1, mRNA :

FOXD3:

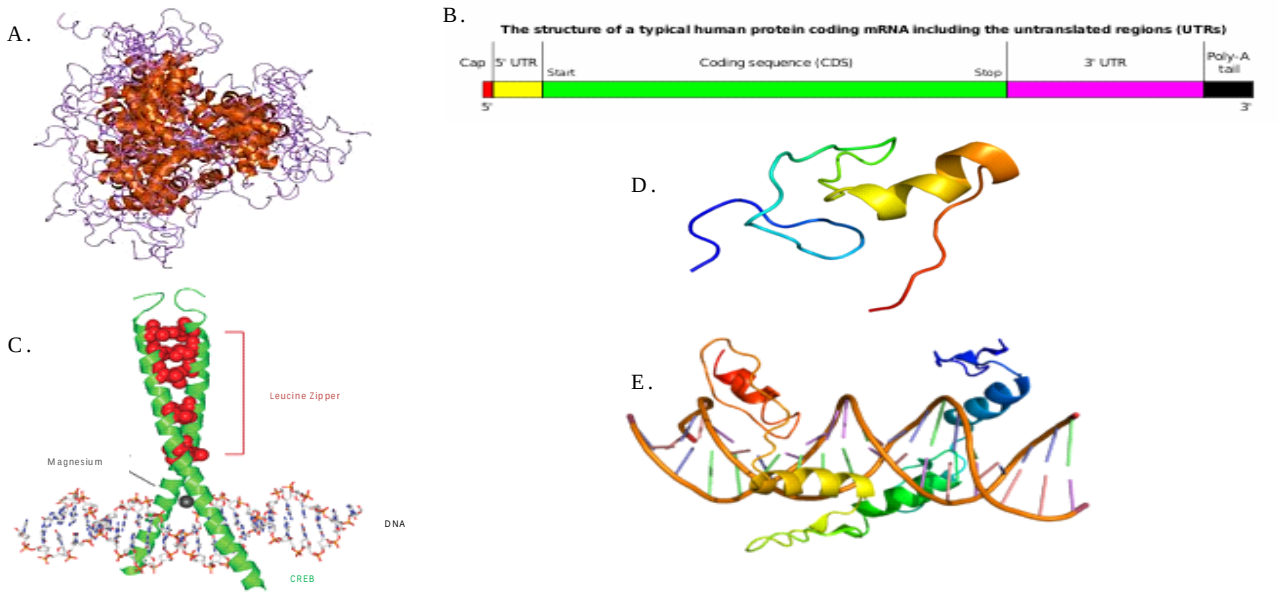
This gene belongs to the forkhead protein family of transcription factors which is characterized by a DNA-binding forkhead domain (Table 7). FOXD3 functions as a transcriptional repressor and contains the C-terminal engrailed homology-1 motif (eh1), which provides an interactive surface with a transcriptional co-repressor Grg4 (Groucho-related gene-4) (Fig. 3A) (Yaklichkin *et al.*, 2007).

Polyadenylation sitesf:

Polyadenylation is the addition of a poly (A) tail to a

Table 7 : UTR characterization is necessary for the studies in the regulation of the gene by RefSeq_RNA database

Sr. No.	Motif type	Motif name	Position	Length	Sequence
1.	polyadenylation sites (PAS)	Polyadenylation sites	128 ~ 159	32	Aaatgatggaggtgattatcttccgttactag
2.	Transcriptional regulatory motif	FOXD3	221 ~ 232	12	Cattgttttatt
3.	Transcriptional regulatory motif	HNF3beta	58 ~ 72	15	Caaattatttgttta
4.	Transcriptional regulatory motif	TATA	206 ~ 220	15	Atataaaagttatgt
5.	Transcriptional regulatory motif	HFH8_(FOXF1A)	64 ~ 76	13	Atttgtttattct
6.	Transcriptional regulatory motif	Msx-1	114 ~ 122	9	Cagtaattg
7.	Transcriptional regulatory motif	IPF1	40 ~ 51	12	Gggttaatgaaa
8.	Transcriptional regulatory motif	AP-3	76 ~ 83	8	Tctaaatt
9.	Transcriptional regulatory motif	LEF1	5 ~ 10	6	Tcaaag
10.	Transcriptional regulatory motif	Gfi1b	48 ~ 59	12	Gaaatcactcca
11.	Transcriptional regulatory motif	HOXA13	208 ~ 213	6	Ataaaa
12.	Transcriptional regulatory motif	DRI1	80 ~ 85	6	Aattaa
13.	Transcriptional regulatory motif	GABP-alpha	147 ~ 152	6	Cttccg
14.	Transcriptional regulatory motif	CRX	1 ~ 6	6	Ctaatc
15.	Transcriptional regulatory motif	MEF-2C	229 ~ 235	7	Tatttat
16.	Transcriptional regulatory motif	PARP	172 ~ 177	6	Tttctt
17.	Transcriptional regulatory motif	Cdx-1	69 ~ 74	6	Tttatt
18.	Transcriptional regulatory motif	Cdx-1	227 ~ 232	6	Tttatt
19.	Transcriptional regulatory motif	CDX-2	201 ~ 206	6	Tttata
20.	Transcriptional regulatory motif	LHX3	81 ~ 86	6	Attaaa
21.	Transcriptional regulatory motif	HOXA5	114 ~ 121	8	Cagtaatt
22.	Transcriptional regulatory motif	Sox4	113 ~ 129	17	Tcagtaattgtttttaa
23.	AU-rich elements (ARE)	AU rich element	226 ~ 238	13	Ttttatttatcat
24.	C-to-U RNA editing sites	RNA C-to-U editing site	-48 ~ -48	1	G



A. Structure of FOXD3; **B**. Polyadenylation sites; **C.** Structure of CREB transcription factor; **D.** GATA 1 transcriptional factor; **E.** YY 1 transcriptional factor t

Fig. 3 : ***In-silico*** **analysis of differentially expressed fragments and their motifs**

messenger RNA. The poly (A) tail consists of multiple adenosine monophosphates; in other words, it is a stretch of RNA that has only adenine bases. In eukaryotes, polyadenylation is part of the process that produces mature messenger RNA (mRNA) for translation. Therefore, it forms part of the larger process of gene expression. The process of polyadenylation begins as the transcription of a gene terminates. The 3'-most segment of the newly made pre-mRNA is first cleaved off by a set of proteins; these proteins then synthesize the poly (A) tail at the RNA's 3' end.

In some genes these proteins add a poly (A) tail at one of several possible sites. Therefore, polyadenylation can produce more than one transcript from a single gene (alternative polyadenylation), similar to alternative splicing (Fig. 3B).

The poly(A) tail is important for the nuclear export, translation, and stability of mRNA. The tail is shortened over time, and, when it is short enough, the mRNA is enzymatically degraded. However, in a few cell types, mRNAs with short poly(A) tails are stored for later activation by re-polyadenylation in the cytosol. mRNA molecules in both prokaryotes and eukaryotes have polyadenylated 3'-ends, with the prokaryotic poly (A) tails generally shorter and less mRNA molecules polyadenylated.

Information about transcriptional motifs found in the >SCoT-21:335:

CREB:

CREB (cAMP response element-binding protein) is a cellular transcription factor (Table 8). It binds to certain DNA sequences called cAMP response elements (CRE), thereby increasing or decreasing the transcription of the downstream genes (Purves *et al*., 2010). It was first described in 1987 as a cAMP-responsive transcription factor regulating the somatostatin gene. It is a transcription factor that regulates diverse cellular responses, including proliferation, survival, and differentiation. CREB is a transcription factor that is known for its role in cell proliferation, differentiation, and survival (Shaywitz *et al.*, 2009; Mayr and Montminy, 2001 and Sakamoto and Frank, 2009). The specific functions of CREB was found in immune responses, including inhibiting NF-kB activation, inducing macrophage survival, and promoting the proliferation, survival, and regulation of T and B lymphocytes (Fig. 3C).

CREB proteins are activated by phosphorylation from various kinases, including PKA and Ca^{2+}/ calmodulin-dependent protein kinases on the Serine 133 residues. When activated, CREB protein recruits other transcriptional coactivators to bind to CRE promoter 5' upstream region. Hydrophobic leucine amino acids are

Table 8:Characterization of >SCoT-21:335 by using Ref Seq_RNA database

Motif type	Motif name	Position	Length	Sequence
Human splicing sites	Acceptor	148 ~ 149	2	Ag
Exon splicing enhancer (ESE)	sc35 - exonic splicing enhancer	137 ~ 144	8	Gactcctt
	sc35 - exonic splicing enhancer	137 ~ 144	8	Gactcctt
	sc35 - exonic splicing enhancer	108 ~ 115	8	Gatccccg
	sc35 - exonic splicing enhancer	108 ~ 115	8	Gatccccg
Exon splicing silencer (ESS)	fibronectin eda exon	222 ~ 227	6	Caagga
Intron splicing enhancer (ISE)	gh-1 intron 3	281 ~ 290	10	Ggtcttggga
	ighg2 cgamma2 (immunoglobulin heavy chain subclass	180 ~ 185	6	Gtgagg
	g2 - cgamma2 gene) - intron 1			
	ctnt, exon 5	315 ~ 321	7	Ggcttga
Transcriptional regulatory motif	ATF	155 ~ 168	14	Atgtgacgtgactt
	CREB	158 ~ 165	8	Tgacgtga
	ATF2:c-Jun	158 ~ 165	8	Tgacgtga
	Ik-1	283 ~ 295	13	Tcttgggaattca
	Ik-2	283 ~ 294	12	Tcttgggaattc
	Ik-3	283 ~ 295	13	Tcttgggaattca
	c-Myc:Max	75 ~ 88	14	Ggctcacgtggaac
	Max	75 ~ 88	14	Ggctcacgtggaac
	CREB	156 ~ 167	12	Tgtgacgtgact
	CREB	156 ~ 167	12	Tgtgacgtgact
	ATF2	156 ~ 167	12	Tgtgacgtgact
	ATF	157 ~ 168	12	Gtgacgtgactt
	aMEF-2	50 ~ 67	18	Atgattaaaaatctccat
	MEF-2	50 ~ 65	16	Atgattaaaaatctcc
	SOX9	82 ~ 95	14	Gtggaacaatgctg
	AREB6	306 ~ 318	13	Aacatacctggct
	E2F	130 ~ 137	8	Tttcgcgg
	E2F	130 ~ 137	8	Tttcgcgg
	E2F	130 ~ 137	8	Tttcgcgg
	E2F-1	130 ~ 137	8	Tttcgcgg
	ATF6	158 ~ 165	8	Tgacgtga
	Arnt	72 ~ 91	20	Tatggctcacgtggaacaat
	PEA3	374 ~ 380	7	Acttcct
	Tel-2	372 ~ 381	10	Tgacttcctg
	ZF5	428 ~ 435	8	Gggcgcca
	USF2	79 ~ 84	6	Cacgtg
	E2F-4:DP-1	130 ~ 137	8	Tttcgcgg
	Ets	370 ~ 381	12	Cttgacttcctg
	Myc	79 ~ 85	7	Cacgtgg
	CREB	155 ~ 165	11	Atgtgacgtga
	CREB,_ATF	157 ~ 165	9	Gtgacgtga
	c-Maf	86 ~ 104	19	Aacaatgctgattgagatt

Table 8 : Contd.……..……..

Table 8 : Contd…….

	c-Ets-2	375 ~ 381	7	Cttcctg
	BEN	35 ~ 42	8	Cagcgcac
	DEC2	76 ~ 85	10	Gctcacgtgg
	ATF-2	155 ~ 166	12	Atgtgacgtgac
	YY1	380 ~ 391	12	Tgcgccatgtgg
	ER81	392 ~ 401	10	Gtcggaaatg
	E2F-3	424 ~ 430	7	Ggcgggg
	LHX3	53 ~ 58	6	Attaaa
	SOX10	86 ~ 92	7	Aacaatg
untranslated region (UTR) motifs	Musashi binding element (MBE)	258 ~ 263	6	Attagt
	Musashi binding element (MBE)	275 ~ 279	5	Gtagt
	Musashi binding element (MBE)	439 ~ 443	5	Gtagt
C-to-U RNA editing sites	RNA C-to-U editing site	-297 ~ -297	1	A
	RNA C-to-U editing site	-70 ~ -70	1	T
	RNA C-to-U editing site	-101 ~ -101	1	C

located along the inner edge of the alpha helix. These leucine residues tightly bind to leucine residues of another CREB protein forming the dimer. This chain of leucine residues forms the leucine zipper motif. The protein also has a magnesium ion that facilitates binding to DNA (Shaywitz *et al*., 1999).

The cAMP response element binding protein (CREB) is a nuclear factor that is regulated by protein kinase A phosphorylation. Transcription is stimulated on binding to the CRE of a phosphorylated CREB dimer, which is held together by leucine zippers. Dimerization and transcriptional efficacy have been found to be stimulated by phosphorylation at several distinct sites, and it has thus, been suggested that CREB may be regulated by multiple kinases. Sequence analysis of the gene has revealed a cluster of protein kinase A, protein kinase C and casein kinase II consensus recognition sites near the N terminus of the protein sequence, and the proximity of these sites to one another indicates the possibility of interaction in a positive or negative fashion to regulate CREB bioactivity (Quinn and Granner, 1990).

Information about transcriptional motifs found in the >SCoT-32:400:

HOXA13:

The HOXA13 gene provides instructions for producing a protein that attaches (binds) to specific regions of DNA and regulates the activity of other genes. On the basis of this role, the HOXA13 gene is called a transcription factor. The HOXA13 gene is part of a larger family of transcription factors called homeobox genes, which act during early embryonic development to control the formation of many body structures (Fujino *et al*., 2002).

GATA1:

GATA1 has been in the floodlight of modern biology as a paradigm for hematopoietic transcription factors in general and GATA factors in particular (Table 9) (Fig. 3D). The GATA family consists of six transcription factors, GATA1 to GATA6. These transcription factors are categorized as a family due to the fact that they all bind to the DNA consensus sequence (A/T)GATA(A/G) by two characteristic C4 (Cys-X2-Cys-X17-Cys-X2-Cys) zinc-finger motifs specific to the GATA family (Ko and Engel, 1993; Martin and Orkin, 1990 and Merika and Orkin, 1993). The DNA-binding regions are highly homologous between the GATA family members. Outside these regions, the conservation between GATA factors is low (Orkin, 1992). The GATA family is divided into two subfamilies on the basis of the expression profiles of the individual transcription factors. GATA1, GATA2 and GATA3 belong to the hematopoietic subfamily, since they are expressed mainly in the hematopoietic system (Weiss and Orkin, 1995). The non-hematopoietic subfamily is composed of GATA4, GATA5, and GATA6, which are expressed in several tissues, including intestine, lung, and heart GATA1, also known as NF-E1, NF-1, Ery-1 and GF-1, is the founding member of the GATA family of transcription factors (Fig. 3D).

Table 9 :Characterization of >SCoT-32:400 by using RefSeq_RNA database

Motif type	Motif name	Position	Length	sequence
Human splicing sites	Donor	325 ~ 326	2	Gt
Exon splicing enhancer (ESE)	sc35 - exonic splicing enhancer	179 ~ 186	8	Gattcgcc
	sc35 - exonic splicing enhancer	179 ~ 186	8	Gattcgcc
	sc35 - exonic splicing enhancer	370 ~ 377	8	Ggtcccta
	sc35 - exonic splicing enhancer	370 ~ 377	8	Ggtcccta
Intron splicing enhancer (ISE)	gh-1 intron 3	409 ~ 418	10	Ggaaggggg
	gh-1 intron 3	271 ~ 280	10	Ggcgctgggc
	gh-1 intron 3	363 ~ 372	10	Ggcggagggt
	gh-1 intron 3	413 ~ 422	10	Gggggggggg
	gh-1 intron 3	414 ~ 423	10	Gggggggggg
	gh-1 intron 3	415 ~ 424	10	Gggggggggg
	gh-1 intron 3	416 ~ 425	10	Ggggggggt
	gh-1 intron 3	420 ~ 429	10	Ggggtgggg
	gh-1 intron 3	421 ~ 430	10	Gggtgggg
	gh-1 intron 3	422 ~ 431	10	Gggtgggggg
	gh-1 intron 3	423 ~ 432	10	Ggtgggggt
	ctnt, exon 5	401 ~ 407	7	Ggctggt
Transcriptional regulatory motif	E2F	268 ~ 275	8	Tttggcgc
	GATA-1	84 ~ 93	10	Cgggatagct
	GATA-2	84 ~ 93	10	Cgggatagct
	Zic1	428 ~ 436	9	Ggggtggta
	MAZR	415 ~ 427	13	Gggggggggtgg
	E2F	268 ~ 279	12	Tttggcgctggg
	Rb:E2F-1:DP-1	268 ~ 275	8	Tttggcgc
	E2F	268 ~ 276	9	Tttggcgct
	E2F	268 ~ 276	9	Tttggcgct
	Churchill	295 ~ 300	6	Cggggg
	Churchill	338 ~ 343	6	Cggggg
	YY1	345 ~ 355	11	Gccgccatatt
	MOVO-B	294 ~ 300	7	Gcggggg
	MOVO-B	337 ~ 343	7	Gcggggg
	SOX10	92 ~ 98	7	Ctttgtc
	Kid3	43 ~ 47	5	Ccacg
	Kid3	81 ~ 85	5	Ccacg
	Kid3	135 ~ 139	5	Ccacg
	GTF2IRD1-isoform2	215 ~ 223	9	Cagattagg
	ZNF515	424 ~ 433	10	Gtgggggtg
	ELF1	408 ~ 413	6	Aggaag
	HOXA13	380 ~ 385	6	Ataaaa
	Pbx1	120 ~ 136	17	Atgctcatcaattgtcc
	SPI1	408 ~ 413	6	Aggaag
	YY1	345 ~ 356	12	Gccgccatatta
	E2F-3	336 ~ 342	7	Ggcgggg
	MITF	322 ~ 328	7	Catgtga
	SOX10	92 ~ 97	6	Ctttgt
Untranslated region (UTR) motifs	K-Box (KB)	323 ~ 330	8	Atgtgata
	Musashi binding element (MBE)	19 ~ 24	6	Gttagt
	Musashi binding element (MBE)	434 ~ 438	5	Gtagt

The GATA factors in soybean exhibited expression diversity among different tissues; some of these factors showed tissue-specific expression patterns. GATA factors displayed upregulation or downregulation in soybean leaf in response to low nitrogen stress.

YY1:

The ubiquitous transcription factor Yin Yang 1 (YY1) is known to have a fundamental role in normal biologic processes such as embryogenesis, differentiation, replication, and cellular proliferation. It exerts effects on genes involved in these processes via its ability to initiate, activate, or repress transcription depending upon the context in which it binds. Mechanisms of action include direct activation or repression, indirect activation or repression via cofactor recruitment, or activation or repression by disruption of binding sites or conformational DNA changes. Its activity is regulated by transcription factors and cytoplasmic proteins that have been shown to abrogate or completely inhibit YY1-mediated activation or repression. Since expression and function of YY1 are known to be intimately associated with progression through phases of the cell cycle, the physiologic significance of YY1 activity has recently been applied to models of tumor biology. It is important to divulge how complex factors such asYY1 function in diverse biological processes and ultimately shape the growth and viability of eukaryotic cells (Gordon *et al*., 2006) (Fig. 3E).

DEF involved in removal of an essential palmitoleate moiety from Wnt proteins :

Protein notum homolog (*Glycine max*):

The DEF derived from SCoT 14-S-2 was found to be up-regulated in symptomatic plant. The *in-silico* studies UniProt link for the most relevant hit revealed identity with *ARABIDOPSIS THALIANA* CHORIS MATE MUTASE-1. It plays an important role in plant growth and developmental stages like flowering, petal differentiation and expansion stage.

Phenylalanine, tyrosine and tryptophan have a dual biosynthetic role in plants; they are required for protein synthesis and are also precursors of number of aromatic secondary metabolites critical to normal development and stress responses. Whereas, much have been learned in recent years about the genetic control of tryptophan biosynthesis in Arabidopsis and other plants. However, relatively little is known about the genetic regulation of phenylalanine and tyrosine synthesis. Each isoform may perform distinct physiological role in co-ordinating chorismatemutaseT activity with developmental and environmental signals.

DEF involved in post-translational modification that regulates cellular pathways :

LEC14B homolog isoform of X1 (*Glycine max*) :

The down regulated DEF from SCoT 14-S-2 was recovered from symptomatic plant. The *in-silico* studies UniProt link for the most relevant hit revealed 14.73 per cent identity. DNA-dependent RNA polymerase catalyzes the transcription of DNA into RNA consuming the four ribonucleoside triphosphates as substrates. Component of RNA polymerases IV and V which regulate accumulation of short-interfering RNAs (siRNA) and subsequent RNA-directed DNA methylation-dependent (RdDM) transcriptional gene silencing (TGS) of endogenous repeated sequences including transposable elements.

Structural localization of differentially expressed fragments by virtual karyotyping :

The soybean genome browsers (Phytozome, SoyBase and NCBI) were used to anchor identified DEF sequences generated on *G.max* virtual chromosomes to identify their distribution in the genome, relative position, and abundance. For this purpose the 'MegaBLAST' tool was used to ascertain the location of DEFs in the genome. The generated virtual karyotype of DEFs sequence produced an anchorage pattern at different locations in chromosomes as depicted in Fig. 4 and details are given in Table 10. This result predicts the different locations on chromosomes by virtual karyotyping with E value ranges between 0.18 to 8.6.

Conclusion:

Our research demonstrated that the The cDNA RAPD, cDNA-SCoTtechniques are easy, less expensive, time-saving, and efficient methods for differential gene expression research, was exploited in this study to identify differentially expressed fragments during floral bud distortion in soybean.In response to bud distortion, we identified 40 differentially expressed DEFs and characterized 5 reproducible DEFs for their homology, protein interactome, sub-cellular localization and physical

Table 10: Showing locations on chromosomes by virtual karyotyping

Primer: Scot 14- As-1

Chr	Map element	Type	Hits	Score	E value
1	NC_016088 *Glycine max* cultivar Williams 82 chromosome 1, V1.1, whole genome shotgun sequence	SEQUENCE	17	37.4	2.2
2	NC_016089 *Glycine max* cultivar Williams 82 chromosome 2, V1.1, whole genome shotgun sequence	SEQUENCE	14	39.2	0.62
3	NC_016090 *Glycine max* cultivar Williams 82 chromosome 3, V1.1, whole genome shotgun sequence	SEQUENCE	17	37.4	2.2
4	NC_016091 *Glycine max* cultivar Williams 82 chromosome 4, V1.1, whole genome shotgun sequence	SEQUENCE	25	35.6	7.5
5	NC_016092 *Glycine max* cultivar Williams 82 chromosome 5, V1.1, whole genome shotgun sequence	SEQUENCE	22	37.4	2.2
6	NC_016093 *Glycine max* cultivar Williams 82 chromosome 6, V1.1, whole genome shotgun sequence	SEQUENCE	9	37.4	2.2
7	NC_016094 *Glycine max* cultivar Williams 82 chromosome 7, V1.1, whole genome shotgun sequence	SEQUENCE	14	37.4	2.2
8	NC_016095 *Glycine max* cultivar Williams 82 chromosome 8, V1.1, whole genome shotgun sequence	SEQUENCE	10	41.0	0.18
9	NC_016096 *Glycine max* cultivar Williams 82 chromosome 9, V1.1, whole genome shotgun sequence	SEQUENCE	5	37.4	2.2
10	NC_016097 *Glycine max* cultivar Williams 82 chromosome 10, V1.1, whole genome shotgun sequence	SEQUENCE	13	37.4	2.2
11	NC_016098 *Glycine max* cultivar Williams 82 chromosome 11, V1.1, whole genome shotgun sequence	SEQUENCE	4	35.6	7.5
12	NC_016099 *Glycine max* cultivar Williams 82 chromosome 12, V1.1, whole genome shotgun sequence	SEQUENCE	7	35.6	7.5
13	NC_016100 *Glycine max* cultivar Williams 82 chromosome 13, V1.1, whole genome shotgun sequence	SEQUENCE	4	35.6	7.5
14	NC_016101 *Glycine max* cultivar Williams 82 chromosome 14, V1.1, whole genome shotgun sequence	SEQUENCE	11	37.4	2.2
15	NC_016102 *Glycine max* cultivar Williams 82 chromosome 15, V1.1, whole genome shotgun sequence	SEQUENCE	15	35.6	7.5
16	NC_016103 *Glycine max* cultivar Williams 82 chromosome 16, V1.1, whole genome shotgun sequence	SEQUENCE	11	37.4	2.2
17	NC_016104 *Glycine max* cultivar Williams 82 chromosome 17, V1.1, whole genome shotgun sequence	SEQUENCE	12	37.4	2.2
18	NC_016105 *Glycine max* cultivar Williams 82 chromosome 18, V1.1, whole genome shotgun sequence	SEQUENCE	21	37.4	2.2
19	NC_016106 *Glycine max* cultivar Williams 82 chromosome 19, V1.1, whole genome shotgun sequence	SEQUENCE	27	37.4	2.2
20	NC_016107 *Glycine max* cultivar Williams 82 chromosome 20, V1.1, whole genome shotgun sequence	SEQUENCE	16	37.4	2.2

Table 10: Contd.……

Table 10 : Contd…….

Primer: Scot 14-S-1					
Chr	Map element	Type	Hits	Score	E value
10	NC_016097 *Glycine max* cultivar Williams 82 chromosome 10, V1.1, whole genome shotgun sequence	SEQUENCE	1	35.6	5.3
11	NC_016098 Glycine max cultivar Williams 82 chromosome 11, V1.1, whole genome shotgun sequence	SEQUENCE	1	37.4	1.5
19	NC_016106 *Glycine max* cultivar Williams 82 chromosome 19, V1.1, whole genome shotgun sequence	SEQUENCE	1	35.6	5.3
Primer:Scot 14–S-2					
Chr	Map element	Type	Hits	Score	E value
3	NC_016090 *Glycine max* cultivar Williams 82 chromosome 3, V1.1, whole genome shotgun sequence	SEQUENCE	1	35.6	8.6
4	NC_016091 *Glycine max* cultivar Williams 82 chromosome 4, V1.1, whole genome shotgun sequence	SEQUENCE	1	35.6	8.6
11	NC_016098 *Glycine max* cultivar Williams 82 chromosome 11, V1.1, whole genome shotgun sequence	SEQUENCE	1	35.6	8.6
12	NC_016099 *Glycine max* cultivar Williams 82 chromosome 12, V1.1, whole genome shotgun sequence	SEQUENCE	1	37.4	2.5
13	NC_016100 *Glycine max* cultivar Williams 82 chromosome 13, V1.1, whole genome shotgun sequence	SEQUENCE	1	35.6	8.6
14	NC_016101 *Glycine max* cultivar Williams 82 chromosome 14, V1.1, whole genome shotgun sequence	SEQUENCE	1	35.6	8.6
20	NC_016107 *Glycine max* cultivar Williams 82 chromosome 20, V1.1, whole genome shotgun sequence	SEQUENCE	1	35.6	8.6
Primer:Scot 15 - S -1					
Chr	Map element	Type	Hits	Score	E value
4	NC_016091 *Glycine max* cultivar Williams 82 chromosome 4, V1.1, whole genome shotgun sequence	SEQUENCE	1	35.6	4.7
13	NC_016100 *Glycine max* cultivar Williams 82 chromosome 13, V1.1, whole genome shotgun sequence	SEQUENCE	1	35.6	4.7
Primer: Scot 13 –S-1					
Chr	Map element	Type	Hits	Score	E value
7	NC_016094 *Glycine max* cultivar Williams 82 chromosome 7, V1.1, whole genome shotgun sequence	SEQUENCE	1	35.6	8.0

karyotyping. The more number of RAPD, SCoT markers will enhance ability to understand molecular aspects in reference to the development of floral bud distortion and specific markers can be developed for screening of plants at early stages. To the best of our knowledge, this study is the first to report differentially expressed fragment in response to floral bud distortion in soybean at molecular level.

Acknowledgement :

We gratefully acknowledge the Board of Research in Nuclear Sciences (BRNS) - Department of Atomic

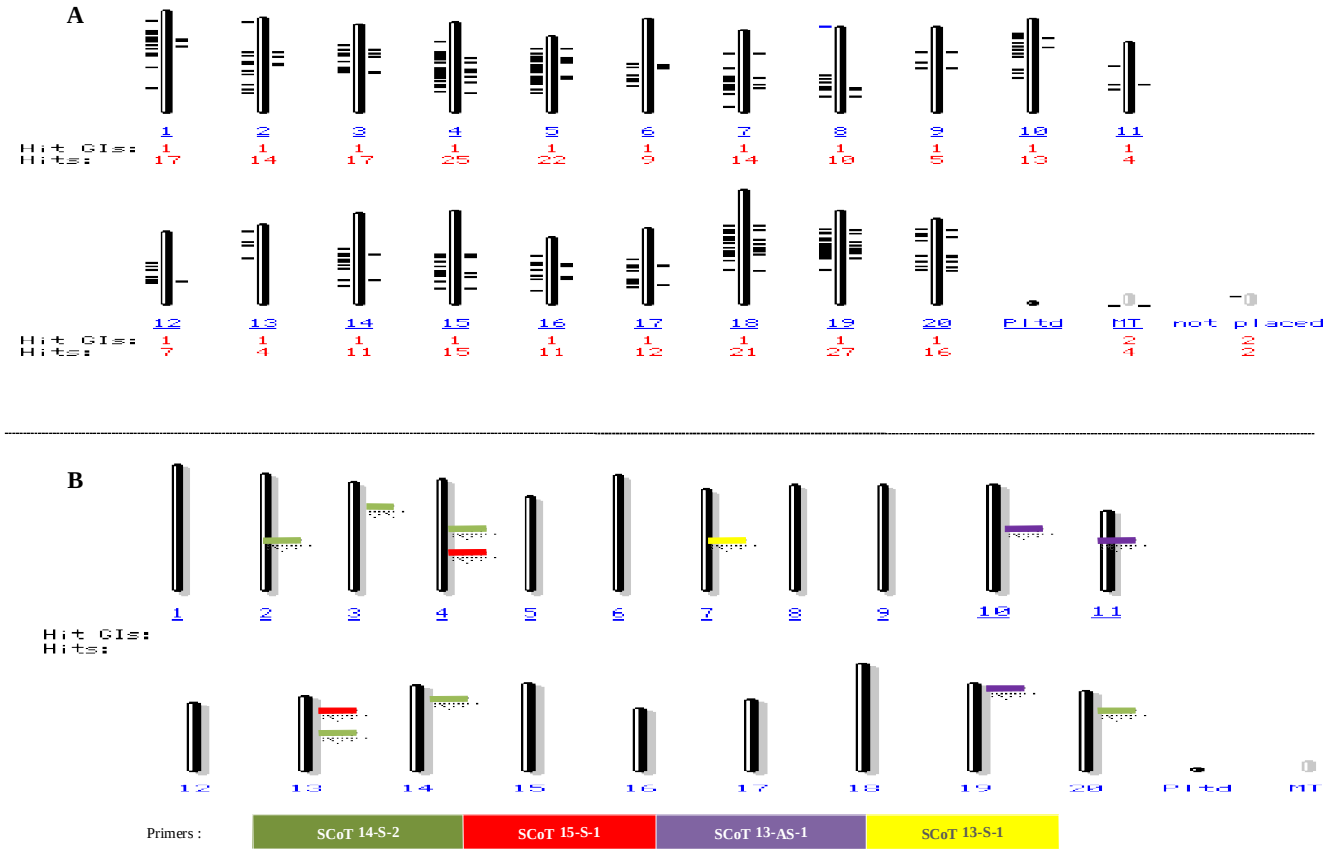


Fig. 4 : ***In-silico*** **karyotyping of sequences (DEFs) on soybean chromosome**

Energy (DAE), Government of India, Trombay, Mumbai, India for providing financial support through a research project (No. 2013/37B/44/BRNS/1904) to carry out this research.

Research project (No. 2013/37B/44/BRNS/1904) to carry out this research.

Authors' affiliations :

S. KAD, M. P. MOHARIL, A. G. DESHMUKH, S. S. MANE AND R. S. NANDANWAR, Post Graduate Institute, Dr. Panjabrao Deshmukh Krishi Vidyapeeth, AKOLA (M.S.) INDIA

P.B. KALE, ICAR-National Research Centre on Plant Biotechnology, Pusa Campus, PUSA (NEW DELHI) INDIA

M.S. DUDHARE, Vasantrao Naik College of Agricultural Biotechnology, Dr Panjabrao Deshmukh Krishi Vidyapeeth, YAVATMAL (M.S.) INDIA

J. G. MANJAYA, Nuclear Agriculture and Biotechnology Division, Bhabha Atomic Research Centre, Trombay, MUMBAI (M.S.) INDIA

REFERENCES

Altschul, S.F., Madden, T.L. and Schaffer, A.A. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, **25**: 3389-3402.

Arunkumar, P. and Chendrayan, K. (2008).Rhizobium inoculants for increasing pulse production.*Biofertilizers Newsletter*, July : 11-13.

Bhatia, V.S., Gupta,G.K., Anita R., Sharma, A.N. and Joshi, O.P. (1999). MLO- nematode suspected in causing no-podding syndrome of soybean.*ICAR News*, **5** (2) : 1-2.

Blackshaw, S., Croix, B.S., Polyak, K., Kim, J.B. and Cai, L. (2007). Serial analysis of gene expression (SAGE): Experimental method and data analysis. *Curr.Protoc. Hum. Genet.Chapter,* **11**: Unit 11.7. John Wiley and Sons, Inc. Maryland.

Bowler, L.D. (2004).Representational difference analysis of cDNA. *Methods Mol. Med.*,**94** (2): 49-66.

Christensen, A.B., Ho, Cho, B., Næsby, M., Gregersen, P.L., Brandt, J., MadrizOrdeñana, K., Collinge, D. and Thordal-Christensen, H. (2002). The molecular characterization of two barley proteins establishes the novel PR-17 family of pathogenesis-related proteins. *Molecular Plant Pathology* **3**, 135-144.

Collard, B.C.Y. and Mackill, D.J. (2009). Start codon targeted

(SCoT) polymorphism: a simple, novel DNA marker technique for generating gene-targeted markers in plants. *Plant Mol. Biol. Report*, **27**: 86–93.

Dafoe, N.J., Gowen, B. E. and Constabel, C. P. (2010).Thaumatin-like proteins are differentially expressed and localized in phloem tissues of hybrid poplar. *BMC Pl. Biol.*, **10**(1): 191.

Ellis, C. M., Nagpal, P., Young J. C., Hagen G., Guilfoyle T. J. and Reed, J. W. (2005). Auxin response factor 1and auxin response factor regulate senescence and floral organ abscission in *Arabidopsis thaliana*. *Development*, **132** : 4563-4574. 10.1242/dev.02012.

Franceschini, A., Szklarczyk, D., Frankild, S., Kuhn, M., Simonovic, M., Roth, A. and Jensen, L.J. (2013). STRING v9. 1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res.,* **41** (D1): D808-D8.

Fujino, T., Suzuki, A., Ito, Y., Ohyashiki, K., Hatano, Y., Miura, I. and Nakamura, T. (2002). Single-translocation and double-chimeric transcripts: detection of NUP98-HOXA9 in myeloid leukemias with HOXA11 or HOXA13 breaks of the chromosomal translocation. *Blood.*, **99**(4) : 1428-1433.

Golnaraghi, A.R., Shahraeen, N., Pourrahim, R., Farzadfar, S. and Ghasemi, A. (2004).Occurrence and relative incidence of viruses infecting soybeans in Iran. *Plant Disease*, **88**:1069-1074.

Gopalan, S., Wei, W. and He, S.Y. (1996).hrp gene-dependent induction of hin1:a plant gene activated rapidly by both harpins and avrPto gene-mediated signal. *Plant J.*, **10**: 591–600.

Gordon, S., Akopyan, G., Garban, H. and Bonavida, B. (2006). Transcription factor YY1: structure, function, and therapeutic implications in cancer biology. *Oncogene*, **25** (8) : 1125-1142.

Gray, W.M., del Pozo, J.C., Walker, L., Hobbie, L., Risseeuw, E., Banks, T., Crosby, W.L. Yang, M. Ma, H. and Estelle, M. (1999). Identification of an SCF ubiquitin-ligase complex required for auxin response in Arabidopsis thaliana, *Genes & Develop.*, **13** : 1678-1691.

Gupta, S., Bharalee, R., Bhorali, P., Das, S.K., Bhagawati, P., Bandyopadhyay, T., Gohain, B., Agarwal, N., Ahmed, P., Borchetia, S. and Kalita, M.C. (2013). Molecular analysis of drought tolerance in tea by cDNA-AFLP based transcript profiling. *Molecular Biotechnol.*, **53**(3) : 237-248.

Hagen, G. and Guilfoyle, T. (2002). Auxin-responsive gene expression: Genes, promoters and regulatory factors. *Plant Mol. Biol.*, **49** : 373–385.

Heim, M.A., Jakoby, M., Werber, M., Martin, C., Weisshaar, B., Bailey, P.C. (2003). The basic helix-loop-helix transcription factor family in plants: A genome-wide study of protein structure and functional diversity. *Mol.Biol. Evol.*, **20**:735.

Hesse, T., Feldwisch, J., Balshuesemann, D. *et al*.(1989). Molecular cloning and structural analysis of a gene from Zea mays (L.) coding for a putative receptor for the plant hormone auxin. *EMBO J.*, **8** :2453-2461.

Hillmann, A., Dunne, E. and Kenny, D. (2009). cDNA amplification by SMART-PCR and suppression subtractive hybridization (SSH)-PCR. *Methods Mol.Biol.*, **49**(6): 223-243.

Horton, P., Park, K.J., Obayashi, T. *et al.* (2007).WoLF PSORT: protein localization predictor. *Nucleic Acids Res.*, **35**: 585–587. doi:10.1093/nar/gkm259.

Jadhav, P.V., Mane, S.S., Nandanwar, R.S. *et al.*(2013). Floral bud distortion in soybean and incidence in Central India. *Egyptian J. Biol.*, **15**(1): 59-65.

Jones, A.M. and Dangl, J.L. (1996). Logjam at the styx: programmed cell death in plants.*Trends Microbiol.*, **1**: 114 - 119.

Ko, L. J. and Engel, J. D. (1993). DNA-binding specificities of the GATA transcription factor family. *Molecular & Cellular Biol.*, **13**(7): 4011-4022.

Koti, S., Reddy, K.R., Reddy, V.R. and Zhao, D. (2005).Interactive effects of carbon dioxide, temperature, and ultraviolet-B radiation on soybean (*Glycine max* L.) flower and pollen morphology, pollen production, germination and tube lengths.*J. Exp. Bot.*, **56** : 725-736.

Lee, M., Bottner-Parker, K. D., Zhao, Y., Villalobos, W. and Moreira, L. (2011). Candidatus phytoplasma costaricanum' a novel phytoplasma associated with an emerging disease in soybean (*Glycine max*). *Int. J. Syst. Evol. Micr.*, **61** : 2822–2826.

Liang, P., Meade, J.D. and Pardee, A.B. (2007). A protocol for differential display of mRNA expression using either fluorescent or radioactive labeling. *Nat Protoc.*, **2**(3): 457-470.

Liu, Y., Zhang, A., Jia, J. and Li, A. (2007).Cloning of salt stress responsive cDNA from wheat and resistant analysis of differential fragment SR07in transgenic tobacco. *J. Genet. Genomics.*, **34** (9) : 842-850.

Lurin, C., Andrés, C., Aubourg, S., Bellaoui, M., Bitton, F., Bruyère, C., Caboche, M., Debast, C., Gualberto, J., Hoffmann, B., Lecharny, A., Le, Ret M., Martin-Magniette, M.L., Mireau, H., Peeters, N., Renou, J.P., Szurek, B., Taconnat, L. and Small, I.(2004). Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. *Plant Cell.*, **16** (8): 2089–2103.

Martin, D. I. and Orkin, S. H. (1990). Transcriptional activation and DNA binding by the erythroid factor GF-1/NF-E1/Eryf 1. *Genes. & Develop.*, **4** (11): 1886-1898.

Mayr, B. and Montminy, M. (2001).Transcriptional regulation

by the phosphorylation-dependent factor CREB. *Nat. Rev. Molecular Cell Biol.,* **2** (8) : 599-609.

Merika, M. E. and Orkin, S. H. (1993).DNA-binding specificity of GATA family transcription factors. *Molecular & Cellular Biol.*, **13**(7): 3999-4010.

Morris, M. D., Moore, L. M. and McKay, M. D. (2006). Sampling plans based on balanced incomplete block designs for evaluating the importance of computer model inputs. *J. Statistical Planning & Inference,* **136** (9) : 3203-3220.

Nimbalkar, S.B., Harsulkar, A.M., Giri, A.P., Sainani, M.N., Franceschi. V. and Gupta, V.S. (2006). Differentially expressed gene transcripts in roots of resistant and susceptible chickpea plant (*Cicer arietinum*) upon *Fusarium oxysporum* infection. *Physiol. Mol. Plant Pathol.*, **668**(4): 176-188.

Nettuwakul, C., Pongtongkam, P., Thongpan, A. and Peyachoknagul, S. (2007). Detection of photoperiod responsive gene in KDML 105 rice (*Oryza sativa* L.) using cDNA-SRAP Technique.*Kasetsart J. Nat. Sci.*, **41** : 651 –659.

Orkin, S.H. (1992). GATA-binding transcription factors in hematopoietic cells. *Blood.*, **80** (3): 575-581.

Polesani, M., Desario, F., Ferrarini, A., Zamboni, A., Pezzotti, M., Kortekamp, A. and Polverari, A. (2008). cDNA-AFLP analysis of plant and pathogen genes expressed in grapevine infected with *Plasmoparaviticola*. *BMC Genomics,* **9**: 142-151.

Pracros, P., Renaudin, J., Eveillard, S., Mouras, A. and Hernould, M. (2006). Tomato flower abnormalities induced by stolburphytoplasma infection are associated with changes of expression of floral development genes. *Mol. Plant Microbe. Interact.*, **19**: 62–68.

Purves, J., Cockayne, A., Moody, P.C. and Morrissey, J.A. (2010). Comparison of the regulation, metabolic functions, and roles in virulence of the glyceraldehyde-3-phosphate dehydrogenase homologues gapA and gapB in *Staphylococcus* aureus. *Infection & Immunity,***78**(12) : 5223-5232.

Quinn, P. G. and Granner, D. K. (1990). Cyclic AMP-dependent protein kinase regulates transcription of the phosphoenol pyruvate carboxykinase gene but not binding of nuclear factors to the cyclic AMP regulatory element. *Molecular & Cellular Biol.*, **10** (7): 3357-3364.

Rahimian, H., Sari H.Z. and Montazerl, M. (1995). Viruses associated with the soybean pod set failure syndrome in Iran. Proceeding of the 12th Iranian Plant Protection Congress 2-7 September 1995, Karadj.

Reyna-Lopez , G.E., Simpson, J. and Ruiz-Herrera, J. (1997). Differences in DNA methylation patterns are detectable during the dimorphic transition of fungi by amplification of restriction Polymorphisms. *Mol. Gen. Genet.*, **253** :703–710.

Rogers, H.J. (2005). Cell death and organ development in plants.*Curr. Topic. Develop. Biol.*, **71**: 225–261.

Sakamoto, K.M. and Frank, D.A. (2009). CREB in the pathophysiology of cancer: implications for targeting transcription factors for cancer therapy. *Clin. Cancer Res.*,**15**: 2583–2587.

Samach, A. and Coupland, G. (2000). Time measurement and the control of flowering in plants.*Bioessays,* **22**: 38–47.

Sarkar, N. (1997). Polyadenylation OFmRNA IN Prokaryotes. *Annl. Rev. Biochem.*, **66**(1) : 173-197.

Shaywitz, A. J. and Greenberg, M.E. (1999). CREB: a stimulus-induced transcription factor activated by a diverse array of extracellular signals. *Annl. Rev. Biochem.,* **68** (1) : 821-861.

Soares-Cavalcanti, N.M., Belarmino, L.C., Kido, E.A., *et al.*(2012). Overall picture of expressed heat shock factors in *Glycine max*, *Lotus japonicus* and *Medicagotruncatula. Genet. Mol.Biol.,* **35** (1): 247-259.

Stintzi, A., Heitz, T., Prasad, V., Wiedemann-Merdinoglu, S., Kauffmann, S.,Geoffroy, P., Legrand, M. and Fritig, B. (1993). Plant "pathogenesis-related" proteins and their role in defense against pathogens. *Biochemie,* **75** : 687-706.

Thomas, S. and Franklin-Tong, V.E. (2004). Self-incompatibility triggers programmed cell death in Papaver pollen. *Nature,* **429**: 305–309.

Toledo-Ortiz, G., Huq, E. and Quail, P.H. (2003). The Arabidopsis basic/ helix-loop-helix transcription factor family. *Plant Cell,* **15**:1749-1770.

Van Loon, L.C., Rep, M. and Pieterse, C.M.J. (2006). Significance of inducible defenserelated proteins in infected plants. *Annual Rev. Phytopathol.*, **44** : 1-28.

Velazhahan, R., Datta, S. K. and Muthukrishnan, S. (1999). The PR-5 family: thaumatin-like proteins. *Pathogenesis-related Proteins in Plants*, 107-129.

Velculescu, V.E., Zhang, L., Vogelstein, B. and Kinzler, K.W. (1995). Serial analysis of gene expression. *Sci.*, **270**: 484-487.

Wee, C.W., Lee, S.F., Robin, C. and Heckel, D.G. (2008). Identification of candidate genes for fenvalerate resistance in *Helicoverpaarmigera*using cDNA-AFLP.*Insect Mol.Biol.*, **17**(4) : 351-360.

Weiss, M. J. and Orkin, S.H. (1995). GATA transcription factors: key regulators of hematopoiesis. *Experim. Hematol.*, **23**(2): 99.

Wu, Y., Dey, R., Han, A., Jayathilaka, N., Philips, M., Ye, J. and Chen, L. (2010). Structure of the MADS-box/MEF2 domain of

MEF2A bound to DNA and its implication for myocardin recruitment. *J. Mol. Biol.,* **397** : 520–533. doi: 10.1016/ j.jmb.2010.01.067.

Xiong, F.Q., Tang, R.H., Chen, Z.L., Pan, L.H. and Zhuang, W.J. (2009). SCoT: a novel gene targeted marker technique based on the translation start codon. *Mol. Plant Breed.,* **7**: 635 – 638.

Xu, Y.H., Yang, F., Tian, X.H., Zhang, W.Y., Wang, X.Y. andMa, G.H. (2011). Different patterns of transcriptomic response to high temperature between diploid and tetraploid *Dioscoreazingiberensis* C. H. *Afr. J. Biotechnol.,* **10**(44): 8847-8854.

Yaklichkin, S., Steiner, A.B., Lu, Q. and Kessler, D.S. (2007). FoxD3 and Grg4 physically interact to repress transcription and induce mesoderm in Xenopus. *J. Biol. Chem.*. **282** (4): 2548–5.

Zamharir, M.G., Mardi, M., Alavi, S.M., *et al*.(2011). Identification of genes differentially expressed during interaction of Mexican lime tree infected with *Candidatus Phytoplasmaaurantifolia*. *BMC Microbiol.,* **11**(1): 2-9.

Zhang, C. and Zhang, F. (2015).The multifunctions of WD40 proteins in genome integrity and cell cycle progression.*J. Genomics.,* **5** (3) : 40-50.

■ WEBLIOGRAPHY

http://www.ncbi.nlm.nih.gov /Blast.cgi.

http://www.ncbi. nlm.nih.gov.