A PREDICTION REVIEW FOR THE INSIGHT TO MOLECULAR MECHANISM UNDERLYING NEGATIVE REGULATION OF *AGAMOUS* (AG) OF *ARABIDOPSIS THALIANA* USING COMPUTATIONAL, BIOCHEMICAL AND MOLECULAR TOOLS

[1]Sumira Malik, [2]Changsoo Kim, [3]Rajnandini Kumari, [4]Sudeep Pathak
[1]Assistant Professor, [2]Professor, [3]Undergraduate student, [4]Undergraduate student
[1]Department of Agriculture and Forestry, Department of Biological sciences and Biotechnology
[1]Tula's institute, Dehradun, India and Chonnam national university, Gwangju, South Korea

Abstract : The floral patterning in *Arabidopsis thaliana* is under regulation of Class A,B,C, E genes. The class C gene *Agamous (AG)* is expressed in whorl 3 and 4 to form stamens and carpel whereas it is repressed in whorl 1 and 2. The previous studies shows the repression of *AG mRNA* through co-repressor LUG, SEU genes but RBE, ANT, SAP and BLR transcription factors are also the repressors of AG through unknown mechanism. This review focuses on repressor's interplay and their role in negative regulation of *AG mRNA* expression in whorl 1 and 2 and so the maturation process of flower development. Here, we target on the preliminary computational approach using Bioconda channel based DREAMTools for the prediction of direct or indirect Protein-Protein interactions and PRD, RAID, PRIDB and RsiteDB tools for Protein- RNA interactions for repression activity of *AG* mRNA through these repressors. Further, confirmatory studies may be performed using biochemical assays such as yeast two hybrid assay, Protein fragment complementation assay, RNA Immunoprecipitation assay and repression assays.

IndexTerms - *Agamous (AG),* DREAMTools, yeast two hybrid assay, PRD, RAID, PRIDB and RsiteDB and biochemical assays

I. INTRODUCTION

The angiosperm flower patterning and development involves several molecular mechanisms. Therefore, the comprehensive and précised understanding into the gene regulatory networks governing this vital developmental process in plants is required. Before two decades, *Arabidopsis thaliana* is one of the intensively investigated flower model for the studies of the molecular mechanism in flower patterning and development. In the recent scenario the molecular mechanisms focuses on interaction among the floral organ A, B, C and E identity genes, specifically the A and C class genes of ABCE model. The combinatorial effect of the A, B, C and E classes of homeotic genes mediates direct the sepals, petals, stamens and carpels development in four concentric whorls, respectively (1, 2, 3). The genes ABCE encode the protein transcription factors also called as MADS box transcription factors. The transcription factors of MADS box possess possibility to interact with each other and in the form multimeric complexes they can activate the vice versa gene (3, 4). AGAMOUS (AG) belongs to Class C gene and regulates the development of the floral organs and it is known as one of the floral organ identity factors and the only key regulator of *Arabidopsis thaliana*. Therefore, in order to reveal the function of AG in determination of organ fate, it is important to give deep insight to genes and mechanisms involved at the downstream of this Class C genes. (5)

Accordingly, *AGAMOUS* (*AG*) gene belongs to C-class and specifies stamen identity in whorl three and promotes carpel development in whorl four. In the whorl one and two, the expression of AG is negatively regulated by AINTEGUMENTA (ANT) that promotes petal identity in whorl two (6). RABBIT EARS (RBE) maintains spatial boundaries in the whorl 2 (7). BELLRINGER (BLR) also represses *AG* in floral meristems and STERILE APETALA (SAP) also a negative regulator AG (8, 9). The well-known two direct interacting plant regulatory proteins LEUNIG and SEUSS (LEU and SEU) transcriptional co-repressors (10) and together represses AG. (11, 12, 13). *AG mRNA* consists of *cis*-regulatory sequences in the unknown putative sites and functions in regulation through binding with repressors. These regulatory sequences in the *AG mRNA* might be the target of these repressor proteins. These repressors bind to putative sites *of mRNA* containing cis-regulatory sequences at unknown positions and negatively regulate the expression of *AG mRNA* (14, 15). In the previous studies, biochemical assays were performed to study protein-protein interaction in yeast, Luc reporter assays and repression assays in plant meristematic tissues were performed (16). Previously, computational analysis with Comprehensive Interaction Databases has been implemented for the exploring the interactions among repressor protein their probable *mRNA* Targets (15). The Protein-protein interactions could be studied using bioconda channel mediated DREAMTool. The DREAMTool is utilised for the exploration of bimolecular protein-protein network Inference of proteins (22-23) and also prediction of transcription factor sequence specificity of unknown

proteins within the genome. This data base determines protein-protein interactions among all the pairwise interactions in network of 47 proteins (24).

The RNA–Protein interactions could be intensively studied through the Protein–RNA interaction database (PRD) analysis. The PRD analysis integrates data sources from the 22 organisms including plants at the gene level. This studies focus on the interaction among proteins and proteins-coding mRNAs. The interaction studies include the relevant information curated from other resources that involves binding sites, protein/RNA motifs and biological functions (17). The resource to such protein-RNA based interactions is RAID. This RAID database works upon the feeded entries and integrates gives computational prediction with the experimental setup. It collects the data from 18 resources and published literatures. It has a wide range of recruited target with 1,208,008 RNA–Protein interactions and contains 60 species covering seven categories that include bacteria, plants and vertebrates (18). Furthermore, the another method of computational analyses to achieve the information of the putative binding sites in the wild and mutant target *mRNA* of multimeric repressor proteins, is the protein–RNA Interface Database (PRIDB) tool. This will provide more information of *the* RNA–Protein interfaces system and also the sequence information of the potential putative target RNA-binding sites for the given proteins. It will be targeting the more clear binding information about the unknown potential RNA–Protein interactions. This database contains structural information for 926 RNA–Protein complexes and includes 9689 protein chains and *2074* RNA chains. In the support of luciferase repression assays, PRIDB will be more for the structures of RNA–Protein complexes and their detailed interfaces to users and will be a detailed informative source to analyse RNA–Protein interactions (19). The another data base suggested is RNA binding site Data Base (RsiteDB) database tool which may support for the domain studies to know the orientation of proteins with respect to target mRNAs (20). RsiteDB would estimate the probability of the interactions among RNA nucleotide bases and protein binding pockets and the RNA–Protein complexes spatial arrangements. Here, we suggest the feasibility of existence of different multimeric protein complexes of *LUG, SEU, RBE, ANT, SAP and BLR* repressor proteins in combinations. These repressors or factors may interact in pair or in coordination with each other or form the multimeric protein complexes. Further, these protein complexes will get recruited to *AG mRNA* and negatively regulate AG through unknown putative sites.

In this review, we suggest and discuss the biochemical and computational methods to identify the protein-protein interactions and their orientation through domain analysis in multimeric protein complex. The binding of the multimeric proteins to the *AG mRNA* through unknown putative binding sites could be examined through wild and mutant type binding assays of the target mRNAs which will further contribute in confirmation of the conserved sequences of the target AG mRNA. The repression assays to investigate the binding and repression of wild and mutant *AG mRNA* through multiple protein complex of *RBE, ANT, SAP and BLR* repressor proteins using luciferase reporter assays might play critical role in yeast cells and plant tissues. The analysis of the putative sequence specific interactions of *LUG, SEU, RBE, ANT, SAP and BLR* proteins with *AG mRNA* will also be required for repression of *AG mRNA. This analysis* will be favoured by computational Data Base programmes as PRD and RAID focusing on Protein-RNA interactions. PRIDB and RsiteDB that are site specific data bases will be utilised to study protein complex and *AG mRNA* interactions. The data obtained from the computational analysis will be used preliminary test to confirm the futher biochemical test .This investigative approach may propose the large number authentic proteins and their role in multimeric complexes as transcription factors or repressors targeting *AG mRNA.*

II. RBE, ANT, SAP, BLR protein-protein interaction through Bioconda mediated DREAMTool analysis with LUG and SEU co-repressor protein further confirmed in Yeast cells and *HEK* cells.

LEUNIG (*LUG*) and *SEUSS* (*SEU* transcription factors) negatively regulates the expression of *Agamous* (*AG*). (11,12). The RBE, ANT, SAP, BLR are the repressors of *AG* mRNA in the floral whorls and meristametic tissues (6-10). The mechanism of the interaction of protein-protein repressors is not explored. Bioconda based DREAMtool predicts the network inference of two different proteins and also determines the transcription factor sequence specificity of proteins within the genome. DREAMtool data base will be determine the protein-protein interactions with the consolidated data for the confirmation through biochemical studies. DREAMtool data based protein interactions are explained in (Table 2.1).

Table 2.1: DREAMTool based prediction of protein-protein interaction-Prediction of LUG, SEU, RBE, ANT, SAP and BLR Protein-Protein interaction using DREAMtool prediction data analysis.

Title, type of data base and synapse ID	Tool and DREAM Nick name	Interaction Studies
Protein-Protein Interaction, Network Inference Implemented 2825374	DREAMTool, D2C2	1. LUG protein with RBE, ANT, SAP and BLR. *2.LUG*+RBE, ANT, SAP and BLR>>>binds to SEU
Protein-Protein Interaction, Network Inference Implemented 2825374	DREAMTool, D2C2	1.SEU protein with RBE, ANT, SAP and BLR. *2.*SEU+RBE, ANT, SAP and BLR>>>binds to LUG

Protein-Protein Interaction, Network Inference Implemented 2825374	DREAMTool, D2C2	LUG + SEU protein with RBE, ANT, SAP and BLR.

Bioconda based DREAMTool analysis for the prediction of possible interactions of LUG and SEU with RBE, ANT, SAP, BLR proteins. LUG protein may interact with any of RBE, ANT, SAP and BLR and then complex binds to SEU. SEU protein may interact with any of RBE, ANT, SAP and BLR and then complex binds to LUG. LUG/ SEU protein complex binds with RBE, ANT, SAP and BLR as explained in Table I.

These all repressors may interact with LUG and SEUSS through direct or indirect protein-protein interactions. This introduces the possibility for the interaction of RBE, ANT, SAP, BLR with either LUG or SEU (Figure 1A) or with both (Figure 1B) and then the repressors gets recruited to AG mRNA. The yeast 2 hybrid bait and prey assay and Pull down assay may be performed to analyze protein-protein interaction as suggested in Figure 1 A and Figure 1B**.** The yeast transformation may be performed with standardized protocol **(10)** and b-galactosidase activity was measured in triplicate using the Galacto Light Plus Kit (Applied Biosystems) and normalized with the OD of the culture.

III. ANT, SAP, BLR proteins in domain based Yeast 2hybrid assays in yeast and Protein Fragment Complementation Assay in *HEK cells*.

Proteins may interact with each-other in different orientations. This gives rise to several possibilities of combinations of full length protein interaction or terminal based interaction as N-N terminal, C-N terminal, N-C terminal and C-C terminals. The RBE, ANT, SAP, BLR proteins may interact with either C/N or N/C terminus of LUG/SEU through N-N terminal, C-N terminal, N-C terminal and C-C terminals as explained in (Table 3.1). These repressor proteins may be cloned in specific orientation in yeast vectors to perform yeast 2 hybrid bait and prey assay (10). Similar strategy for cloning may be performed in HEK (Hamster embryonic kidney) cells to analyse domain based protein-protein interactions for Protein Fragment Complementation Analysis (FCA) through mentioned protocol **(**22).

Figure 1. Mode of interaction of LUG and SEU with RBE, ANT, SAP, BLR proteins: The different forms where RBE, ANT, SAP, BLR proteins bind to LUG and SEU as bridge or RBE, ANT, SAP, BLR binds to LEU /SEU independently.

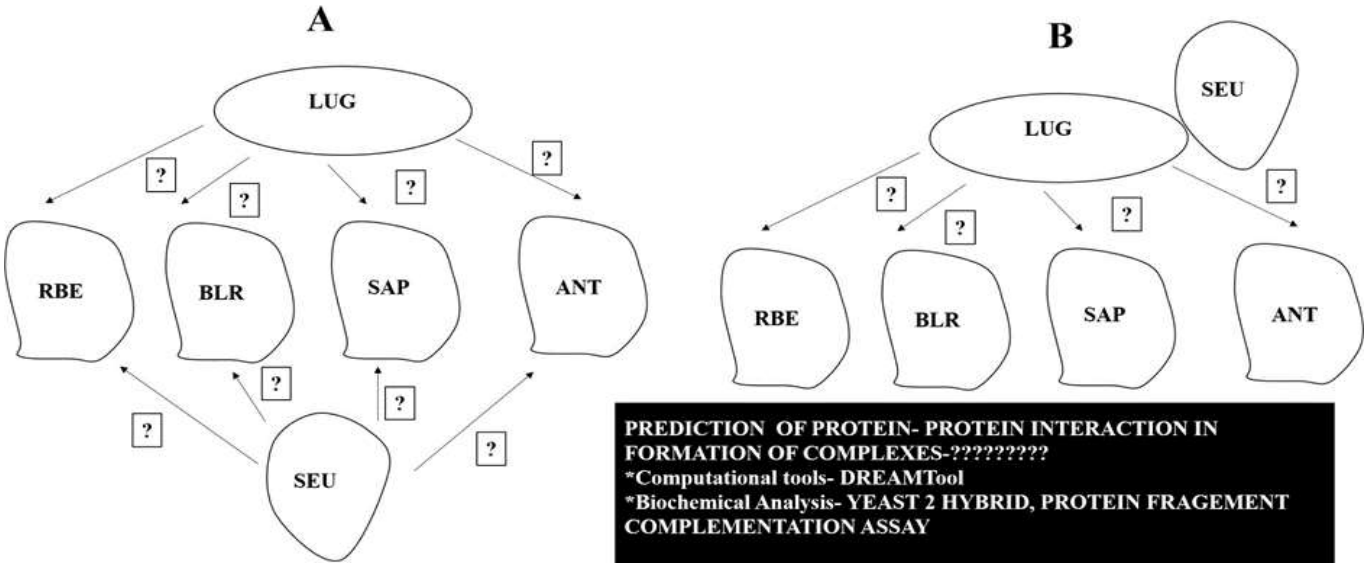


Table 3.1: Protein-protein orientation based domain interaction Test- Orientation of LUG and SEU with RBE, ANT, SAP, BLR proteins in PFCA. A LUG and SEU with RBE, ANT, SAP, BLR proteins interaction (C/N) in various possible combinations for yeast 2 hybrid and PFCA in HEK cells.

Protein A Domain	Protein B domain	Interaction Studies method
LUG or SEU N/C	RBE, ANT, SAP and BLR C/N	PFCA
	RBE, ANT, SAP and BLR N/C	PFCA
	CONTROL	PFCA

IV. The proposed mechanism to detect repression of *AG mRNA* through its putative sites by LUG/ SEU with RBE, ANT, SAP, BLR proteins for the region based screening of wild and mutant screening *AG mRNA*

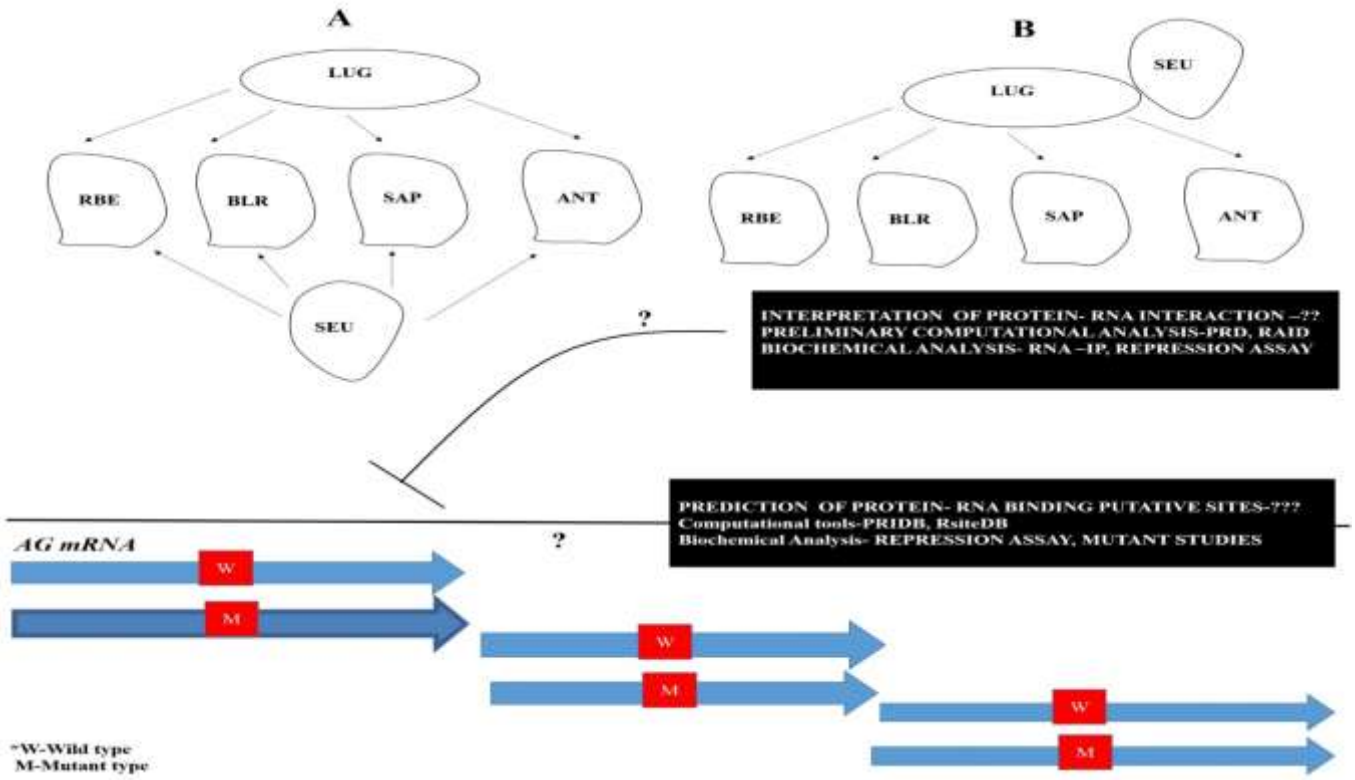
The probable complex of LUG/SEU, RBE, ANT, SAP and BLR repressor proteins with AG mRNA complexes could be predicted preliminarily through Protein–RNA interaction database (PRD) and RNA-associated interaction database (RAID). The detailed studies of the target putative sites in wild and mutant *AG* mRNA could further be dissected and screened according to the region and size of the 3'UTR of *AG* mRNA. The binding sites in the 3'UTR of *AG mRNA* with recruited protein interface in the complex can be revealed through RsiteDB and PRIDB tools respectively. The details are explained in the table. (Table 4.1).

Table 4.1: Information of computational databases- The prediction of Protein-RNA Interactions and binding sites of *AG mRNA*

Name	Abbreviation, version and website	Type of database	INTERACTION STUDIES
Protein–RNA interaction database	PRD (July 2012) http://pri.hgc.jp/	Protein-RNA Interaction	Probable LUG/SEU, RBE, ANT, SAP and BLR repressor proteins complex with *AG mRNA*
RNA-associated interaction database	RAID RAID v2.0 (October 2016)	Protein-RNA Interaction	Probable LUG/SEU, RBE, ANT, SAP and BLR repressor proteins complex with *AG mRNA*
Protein–RNA Interface Database	PRIDB PRIDB v2.0 (October 2010)	Binding Sites Databases.	Probable LUG/SEU, RBE, ANT, SAP and BLR repressor proteins complex with wild and mutant putative sites *AG mRNA*
RsiteDB	RsiteDB RsiteDB (October 2008	Binding Sites Databases.	Probable LUG/SEU, RBE, ANT, SAP and BLR repressor proteins complex with wild and mutant putative sites *AG mRNA*

The overview of databases used for the binding of co-repressors LUG/SEU with RBE, ANT, SAP, BLR repressor proteins at *AG mRNA* might be studied using PRD and RAID. The PRIDB and RsiteDB prediction database can screen and perform acquisition of putative sites in wild type AG mRNA repression through multiple protein complex of LUG/SEU with RBE, ANT, SAP, BLR repressors. After the confirmation from the preliminary computational analysis, the predicted region will be further screened, cloned and intensively studied for the repression of *AG mRNA* by LUG/SEU, RBE, ANT, SAP and BLR repressor proteins through luciferase reporter assay. The putative sites in wild type mRNA will be mutated and luciferase assays which has high repression by protein complex of repressor can be assayed using Dual Luciferase Reporter Assay System (Promega) (10). Further mechanism of screening of putative binding sites and the repression of AG mRNA through probable multiple protein complexes of LUG/SEU, RBE, ANT, SAP and BLR repressor proteins are explained in Figure. 2

Figure 2. Mechanism of repression of *AG* mRNA by LUG/SEU and RBE, ANT, SAP and BLR repressor through screening of different regions - The wild and mutant studies of repression *AG mRNA* confirm its importance in binding and repression through LUG/SEU with RBE, ANT, SAP, BLR factors. The region responsible for repression by this complex could be checked.



V. CONCLUSION:

Previous studies showed the individual and independent involvement of these repressor factors RBE, ANT, SAP and BLR with LUG/SEU in repression of *AG mRNA*. Thus, lots of progress in attaining the knowledge through biochemical, molecular and computational experiments is required to understand the basics of mechanisms that control flower formation of *Arabidopsis thaliana*. The preliminary predictions of protein-protein interactions using DREAMTool a data base for prediction of interface of two different proteins could be used and further the biochemical studies will be performed on the bases of preliminary data through this tool. On the bases of preliminary computational analysis and confirmatory biochemical analysis, protein-protein domain studies of LUG/SEU with RBE, ANT, SAP, BLR through their C or N terminal in Yeast and HEK cells may confirm the orientation and presence of probable multiple protein complex. The multiple protein complex of LUG/SEU with RBE, ANT, SAP, BLR repressor proteins may be involved in promoting the repression of *AG mRNA* in whorl 1 and 2 of *A. thaliana*. The domain analysis might ensure the recruitment of any of the two RBE, ANT, SAP, BLR repressor or the combinatorial effect of most of these repressor proteins with LUG/SEU to *AG mRNA*. Currently, the protein-RNA interaction database and protein specific RNA binding sites databases from PRD, RAID, PRIDB and RsiteDB are one of the essential and accurate sources to explore post-transcriptional regulatory network. The PRD and RAID databases highlight the significant probable protein-RNA interactions and provide suitable platform for the preliminary analysis for the studies of post transcriptional regulatory mechanism in *A, thaliana.* Similarly, the repression of *AG mRNA* through RBE, ANT, SAP, BLR, LUG and SEU altogether or independently through the putative unknown sites in the wild type *AG mRNA* could be studied using PRIDB and RsiteDB that predicts a binding site database of computational analysis.

In this review, we have summarized protein-protein and protein-RNA approach to study the molecular mechanism of repression of *AG mRNA* via repressor proteins or the floral organ identity factors in form of multiple protein complex. Study of these protein-protein and protein-mRNA regulatory mechanisms elucidates the mechanism involved in regulation of molecular mechanisms of Arabidopsis flower formation. In future, it is important to investigate the contribution of floral organ identity transcription factors to their respective mRNAs with more detailed methods of computational studies. The problems with biochemical studies system revealed that cells cannot be readily tracked over an extended developmental time period. Therefore, computational database study might serve as an option for the prediction of protein-protein interaction, prediction of putative binding sites in mRNA and targets of repressor proteins in *Agamous mRNA* in the process of flower development

REFERENCE

[1] Jack, T. (2004). Molecular and genetic mechanisms of floral control. *Plant Cell* 16, S1-S17
[2] Coen, E. S. and Meyerowitz, E. M. (1991). The war of the whorls: genetic interactions controlling flower development.*Nature* 353,31 -37.
[3] Theissen, G. and Saedler, H. (2001). Plant biology. Floral quartets. *Nature* 409,469 -471.
[4] Honma, T. and Goto, K. (2001). Complexes of MADS-box proteins are sufficient to convert leaves into floral organs.*Nature* 409,525 -529

[5] Diarmuid S. Ó'Maoiléidigh, Samuel E. Wuest,LiinaRae, Andre aRaganelli,Patrick T. Ryan,Kamila Kwaśniewska, PradeepDas, Amanda J. Lohan,[c] Brendan Loftus,[c] Emmanuelle Graciet,and Frank Wellmer. (2013) Control of Reproductive Floral OrganIdentity Specification in *Arabidopsis* by the C Function Regulator AGAMOUS. Plant Cell. 2013 Jul; 25 (7): 2482–2503
[6] Krizek BA, Prost V, Macias A (2000) AINTEGUMENTA promotes petal identity andacts as a negative regulator of AGAMOUS. Plant Cell 12: 1357-1366..ANT
[7] Krizek BA, Lewis MW, Fletcher JC. (2006) RABBIT EARS is a second-whorl repressor of AGAMOUS that maintains spatial boundaries in Arabidopsis flowers. Plant J 45: 369-383.
[8] Bao X, Franks RG, Levin JZ, Liu Z. (2004) Repression of AGAMOUS byBELLRINGER in floral and inflorescence meristems. Plant Cell 16: 1478-1489
[9] Byzova MV, Franken J, Aarts MG, de Almeida-Engler J, Engler G, et al. (1999) Arabidopsis, a multifunctional gene regulating inflorescence, flower, and ovule development. Genes Dev 13: 1002-1014.
[10] Sridhar VV, Surendrarao A, Gonzalez D, Conlan RS, Liu Z. (2004) Transcriptional repression of target genes by LEUNIG and SEUSS, two interacting regulatory proteins for Arabidopsis flower development. Proc Natl Acad Sci U S A 101:11494-11499.
[11] Franks RG, Wang C, Levin JZ, Liu Z. (2002) SEUSS, a member of a novel family of plant regulatory proteins, represses floral homeotic gene expression with LEUNIG. Development 129: 253-263.
[12] Liu Z, Meyerowitz EM. (1995) LEUNIG regulates AGAMOUS expression in Arabidopsis flowers. Development 121: 975-991.
[13] Conner J, Liu Z. (2000) LEUNIG, a putative transcriptional corepressor that regulates AGAMOUS expression during flower development. Proc Natl Acad Sci U S A 97: 12902-12907.
[14] Busch, M.A., Bomblies, K., and Weigel, D. (1999). Activation of a floral homeotic gene in*Arabidopsis*. Science 285, 585–587
[15] Deyholos, M. K., & Sieburth, L. E. (2000). Separable whorl-specific expression and negative regulation by enhancer elements within the AGAMOUS second intron. *Plant Cell*, *12*(10), 1799-1810.
[16] Vaniyambadi V. Sridhar, Anandkumar Surendrarao, Zhongchi Liu. (2006). *APETALA1* and *SEPALLATA3* interact with *SEUSS* to mediate transcription repression during flower development.: 3159-3166. Development 133
[17] Fujimori, S.; Hino, K.; Saito, A.; Miyano, S.; Miyamoto-Sato, E. PRD. (2012).A Protein–RNA interaction database. Bioinformation, 8, 729–730.
[18] Yi, Y.; Zhao, Y.; Li, C.; Zhang, L.; Huang, H.; Li, Y.; Liu, L.; Hou, P.; Cui, T.; Tan, P.; et al. RAID v2.0 (2016). An updated resource of RNA-associated interactions across organisms. Nucleic Acids Res., 45,
[19] Lewis, B.A.;Walia, R.R.; Terribilini, M.; Ferguson, J.; Zheng, C.; Honavar, V.; Dobbs, D. Pridb.(2011). A Protein–RNA interface database. Nucleic Acids Res. 39, D277–D282.
[20] Shulman-Peleg, A.; Nussinov, R.; Wolfson, H.J. RsiteDB: (2009), A database of protein binding pockets that interact with RNA nucleotide bases. Nucleic Acids Res 37, D369–D373
[21]. Sumira Malik,Wijeong Jang and Changsoo Kim. (2017) Protein Interaction Mapping of Translational Regulators Affecting Expression of the Critical Stem Cell Factor Nos. Dev Reprod. 21(4): 449–456.
[22] Prill RJ, Marbach D, Saez-Rodriguez J, *et al.*: Towards a rigorous assessment of systems biology models: the DREAM3 challenges. *PLoS One.* 2010; 5(2):e9202.
[23] Weirauch MT, Cote A, Norel R, *et al.*: Evaluation of methods for modeling transcription factor sequence specificity. *Nat Biotechnol.* 2013; 31(2): 126–34.[24] Stolovitzky G, Prill RJ, Califano A: Lessons from the DREAM2 Challenges. *Ann N Y Acad Sci.* 2009; 1158: 159–95.