ISSN: 0974-8571 Vol.13 No. 1 June, 2021

International Journal of Computational Intelligence in Control

Identification of Evolutionary Conserved Promoter-Based Regulatory Sequences for *Senescence-Associated 29* Gene

Zainab[1], Mamoona Rauf[1]*, Humaira Gul[1], Zahoor Ahmad Sajid[2], Muhammad Abid[1], Umair Jan[1], Asim Ur Rahman[1], Behram[1], Izhar Ullah[1], Aizaz Khan[1], Javeria[1], Khayyam Aziz[1], Hussan Ara Begum[1]

[1]Department of Botany, Garden Campus, Abdul Wali Khan University Mardan, 23200. Pakistan

[2]Institute of Botany, University of the Punjab, Lahore, Pakistan

Abstract: Identification of orthologues is very crucial in the fields of comparative genomics for the characterization and exploitation of functionally preserved gene products in crops. *SAG29/SWEET15* a senescence-associated gene is involved in nutrient re-mobilization upon aging. *SAG29* transcript abundance is higher not only in leaves upon abiotic stresses and seeds, showing its crucial involvement to regulate sink remobilization. Current study is carried out to elucidate the gene regulatory network of *SAG29* and its orthologues using *in silico* approach. To this end, a comparative genome-wide bioinformatic analysis was performed for the identification of evolutionarily conserved non-coding sequences (CNSs) in the 1000-bp promoter region (counted from the translation initiation codon). Consensus sequence revealed a highly conserved motif at -673, *Zea may* at -720, and *Glycine max at* -517 upstream from *ATG*. The consensus sequence of this novel putative *cis*-regulatory element is designated from the sequence logo as "NAA(T/A)[ATA]NN(C/G)NNGNAA(T/A)AA" with "ATA" core binding site for upstream transcription factors controlling the expression of *SAG29*. These newly discovered putative *cis*-regulatory elements might have the conserved role in controlling *SAG29* and its orthologues biological role upon stress and development of plants. Moreover, SAG29 expression is also known to be induced in roots of *Glycine max* (harboring *SAG29* orthologue), current research also revealed the negative growth effect induced by sodium selenite, in terms of growth retardation, reduced chlorophyll production, and declined endogenous IAA, indicating the growth inhibitory effect the selenium supplementation in *Glycine max.* Therefore, *SAG29* is a good candidate for the generation of abiotic stress-tolerant *Glycin max.*

Keywords: *Arabidopsis thaliana*, selenium, orthologous, nutrient remobilization, conserved regulatory sequence, promoter, *SAG29*

INTRODUCTION

Selenium (Se) an oxygen group five element, enhanced vegetal bulk each below non-stressed and stressed environments when applied exogenously (Chu *et al.* 2010). In *Arabidopsis* putative-selenium overexpression binding macromolecule increased sodium selenite endurance (Agalou, Roussis and Spaink 2005). Systematically, abscisic acid performs a pivotal role in responses to stress in environment and, hence, longevity of leaf. Abscisic acid is believed to push a leaf's abscission and aging, and the plant hormone levels increase exponentially within the leaves undergoing senescence (Lim *et al.* 2007). Last stage of growth in leaf is leaf senescence cause by speedy chlorophyll loss (Hortensteiner 2006), photochemical efficiency reduction, membrane ion out flow increase senescence-associated genes activation like: *SAG12, SAG13, SAG29, SAG113* (Lohman *et al.* 1994), and photosynthetic genes repression like *RIBULOSE BISPHOSPHATE CARBOXYLASE CHAIN [RBCS]* and chlorophyll *A/B BINDING PROTEIN1 [CAB1]* (Weaver *et al.* 1998). *SAG*s (senescence associated genes) control leaf senescence. *SAG*s are known by transcriptome analysis and many *SAG*s by genetic analysis were found to have an effect on leaf senescence process (Zhonghai Li 2012). Yet one gene cannot be approved as reliable for senescence, the procedures of senescence are evidently in control of

genetic (Nam 1997). Several developments had been got via the classification and characterization of many *SAG*s and mutants related to senescence that offer higher awareness of leaf senescence at the level of molecules (Buchanan-Wollaston *et al.* 2003). Senescing tissues carried the *Arabidopsis SAG29* putatively coding a protein with 2 saliva/MtN3 domains. The copies rise slowly throughout biological leaf senescence, aroused by diffusion stresses through an abscisic acid dependent path. The *Arabidopsis SAG29* is additionally stimulated by low temperature, high level of salinity, and dearth (Seo *et al.* 2011). Stress tolerance affected by responses involve role of the *SAG29* that are involved with results of salinity stress (Balazadeh *et al.* 2010). *SAG29* transgenic overexpression results in increased sensitivity to salt and also premature senescence whereas a loss-of-function alteration in *SAG29* is rarer in response to salt competed with the wild-type (Seo *et al.* 2011).Beside this *SAG29* is present in early seeds, showing that *SAG29* might regulate sink intensity. Thus, the interference of *SAG29* with sink-source interactions will explained the first -senescence constitution of *SAG29* overexpression plants (Chen *et al.* 2015). Gene expression in *SAG29* is evoked by abiotic stresses like low temperature, NaCl, dearth, and ABA treatment (Buchanan *et al.* 2005). *SAG29* protein from *Arabidopsis* could be a basic molecule with medium molecular weight that is associated with completely different members of MtN3_slv protein family. This protein is additionally necessary within the regulation of cell viability and response to salt stress (Peña 2015). Cell viability reduced within the roots under regular development attributes in the transgenic plants 35S:*SWEET15*. In distinction, viability of the cell within the *SAG29*-deficient roots with mutation was not recognizable from the one within the roots of control plants. Arguably, the mutant roots entailed increased viability in cell below high salinity. Overexpressing transgenic *SAG29* plants entailed increasing levels of senescence (Seo *et al.* 2011). Additionally, their protein product are known and functionally characterised in varied plant species (Lim *et al.* 2007). Among the main *SAG* proteins, the proteins MtN3 / saliva / *SWEET* stand out, that possesses seven transmembrane potential motifs and actively participates within the embryonic development and gene activation. The bioinformatics analysis of the protein *SAG29* of *Arabidopsis* attached the response to saline stress in order to explore the molecular mechanisms underlying its function (Meng Yuan 2013). Seventy eight *SAG*s from one hundred seventy downstream genes of ORE1 were majorly up-regulated in plants overexpressing *ORE1* (Balazadeh *et al.* 2010). Direct goals of *ORE1* are *SINA1*, *SWEET15* and *BFN1 (BIFUNCTIONAL NUCLEASE1*) (Matallana *et al.* 2013). ORE1 up-regulated many *NAC* genes via complete leaf senescence (Breeze *et al.*, 2011) was rumored that inducible over-appearance of *GmSARK* (*SENESCENCE-ASSOCIATED RECEPTOR-LIKE kinase in Glycine max*) or *AtSARK (SENESCENCE-ASSOCIATED RECEPTOR-LIKE kinase* in *Arabidopsis*), is the cause of precocious senescence, on the other hand, a T-DNA addition transmuted of *SENESCENCE-ASSOCIATED RECEPTOR-LIKE kinase* in *Arabidopsis* entailed considerably postponed senescence (Xu *et al.* 2011).

In contemporary studies, we will identify and analyze with comprehension the *SAG29* family and their orthologues over the diverse family of species. The work involves the documentation of *SAG29* orthologous gene, a phylogenetic relationship among them and conserved non-coding regulatory motifs in their promoters. By using the accessible expression information in gene ventilator for *SAG29* orthologous gene. A complete study of tissue specific expression in different plant species and their co-expression network were also perform. The co-expression network of the *SAG29* and its orthologous would lead us to find out the homology at the protein level. The evolutionary analysis of cis-regulatory sequences by phylogenetic method will help us to discover the evolutionarily conserved motif present in the promoter of *SAG29* from *Arabidopsis* and its orthologous distant species, thus helpful to understand the expressional activity of this gene.

2: MATERIALS AND METHODS

2.1Bioinformatics characterization of *SAG29 (AT5G13170)*

2.1.1 Selection of *SAG29*

The *Arabidopsis* Information Resource (TAIR) (*https://www.arabidopsis.org*) was used to select the reference gene *SAG29* of *Arabidopsis thaliana* for *elucidating gene regulatory network*, molecular biology and genetics for the model plant *Arabidopsis thaliana* database are available on it. Database of TAIR contains the whole genome sequences of the model plant along with, expression of gene, gene structure ,gene design info, genome maps, seed stocks and DNA, physical markers and genetic, publications and all the *Arabidopsis* research information going on in the scientific community worldwide (Swarbreck *et al.* 2008).

2.1.2 Expression analysis of the *SAG29*

In order to carry out the expression analysis of the *SAG29*, a well-known search engine was used called as "Genevestigator" (*https://www.genevestigator.com*). Genevestigator is a search engine for gene expression which integrates well described public microarray and RNAseq based experiments and attractive visualizations of gene expression covering different biological contexts. (Zimmermann *et al.* 2004).

Gene were analyzed at anatomical and developmental level by recovering the standards of expression from affymetrix array database from Genevestigator. For this, microarray expression data was obtained using *Arabidopsis* Gene Chip platform. For *SAG29* and its selected orthologues gene identifiers remained used as query sequences and carry out explorations in the Gene Chip platform of Genevestigator, where microarray data was analyzed only for the wild type background.

2.1.3 Genome-wide Identification for Orthologues of *SAG29 (AT5G13170)*

To search orthologues of *SAG29*, using nucleotide sequence of the selected gene i.e. *AT5G13170* as query. TAIR (*www.arabidopsis.org*) were used to downloaded the reference genes , which were used as query to execute numerous database searches against the genome data with the help of a well-known tool called as "Phytozome" (www.phytozome.net) (Goodstein *et al.* 2012). While for the second approach, the known gene sequences of *SAG29* rom *Arabidopsis* was used to perform BLAST from the "NCBI" database (*blast.ncbi.nlm.nih.gov*) (Geer *et al.* 2009) to confirm the orthologous genes listed in their respective tables.

For the reference gene *SAG29* nineteen orthologues were selected from Amino acid Identity 79.5% to 55.8% as shown in (**Table 4.1**).

2.1.4 Retrieval of promoter, introns and CDS sequences of the Reference gene and its Orthologues

The database Phytozome was used in order to retrieve the 1000 bp promoter and intron region and CDS sequences of reference genes i.e. *AT5G13170* and its selected orthologues for further evolutionary analyses.

2.1.5 Verification of Orthologous genes on the basis of Evolutionary and phylogenetic Trends

Ninteen closest members of orthologous genes of the reference gene from different species were selected to extract the putative 1000-bp promoter region counted from transcription initiation codon (ATG) and CDSs. The gene annotations and 1-kb promoter regions from various plant species were extracted from NCBI and analyzed the phylogenetic relationship between *SAG29* orthologous gene. The phylogenetic trees were constructed using a multiple sequence alignment tool (MUSCLE) (*www.ebi.ac.uk /muscle*) (Edger 2004). The phylograms were visualized through "TreeDyn" (*www.phylogeny.fr*) (Dereeper *et al.* 2008). High bootstrap values supported by the phylogram helped out in identification of numerous orthologous genes. Maximum likelihood bootstrap consensus phylogeny of CDS and 1kb Promoter sequences of *SAG29* and its orthologous genes. Neighbor-joining distance tree based with bootstrap support is given along the branches.

2.1.6 Screening of the reference genes and its Orthologues on the basis of Expression Pattern Homolgy

To assess the selected (*Glycine max and Zea mays*) of the reference genes on the basis of expression pattern homology at different developmental stages and threshold. For this purpose The Bio-Analytic Resource for Plant Biology "eFP Browser" (*http://www.barutoronto.ca/*) (Winter *et al.* 2007) was used.

2.1.7 Identification of the known *cis*-regulatory elements in 1000 bp Sequences of Reference gene and its Orthologues

Search for *cis*- regulatory elements in the 1kb promoter and introns regions of the reference gene and its orthologues a famous database known as "PlantCARE" was explored (*http://sphinx.rug.ac.be:8080/PlantCARE/*) (Zeng ,Roslinsky and Cheng 2017).. In order to search for plant *cis*-acting regulatory elements out of query sequence in this database, regulatory elements, enhancers and repressors, are represented by consensus sequences, individual sites and positional matrices on specific promoter and introns sequences (Lescot *et al.*, 2002).

2.1.8 Identification of Evolutionarily Conserved Sequences within the promoter of Reference gene and its orthologues

Promoter lies on the upstream of the coding section of all gene, essential for transcription to take place and for enrolment of RNA polymerase. Regulatory regions can also present in the introns of genes and are not restricted to the promoter (Schauer *et al.* 2009). With the help of the "EARS" tool (*wsbc.warwick.ac.uk*) a comparative genome-analysis was performed by which the 1kb promoter and introns sequences of Reference gene and its orthologous genes were analyzed the EARS software checked and analyzed each sequence by degrading them into smaller sub sequences or windows then carried out global alignment between each pair, thus allowed the revealing of conserved sequences (Picot *et al.* 2010). For running this analysis, a cut off P-value of 0.05 and 100bp windows size were taken for promotor and a windows size of 100bp and a cut off P-value of 0.001 were used for introns. The EARS results file for individually run was also separately investigated by software and significant peaks location was identified in the Reference gene promoters and introns of *Arabidopsis* and its orthologous promoter and introns sequences.

2.1.9 Over-representative analysis for putative *cis*-elements in conserved non-coding regions

In order to indicate if the window has a match which is highly conserved in the other species we identified the conserved motif sequence at the region of promoters and introns for

which "MEME suit" (*http://meme.nbcr.net*) (Bailey *et al.* 2009) was used to reveal the consensus sequences between positions of −200 to −1000 of promoter and −200 to −1000 of introns.

2.1.10 Co-Expression Network Identification of Reference and Orthologous genes

In order to reveal the functional interactions of the proteins of Reference and its orthologous genes an online resource and STRING (a search tool database for the retrieval of interacting gene/protein) *(http://string-db.org/)* was accessed. It offers a unique and wide-range coverage and easy access to both predicted along with experimental networking info. interactive network viewer, STRING can homology models, updated previews of structural information and cluster networks widespread data informs and intensely upgraded combination with the third event properties (Szklarczyk *et al.* 2011).

2.1.11 Workflow for *de-novo* Identification/Prediction of putative *cis*-regulatory elements

Stepwise workflow for *de-novo* Identification/Prediction of putative *cis*-regulatory elements (pCREs) in the 1kb promoter sequences of *SAG29* and their orthologues in selected plant species, is mentioned in **(Fig 2.1).**

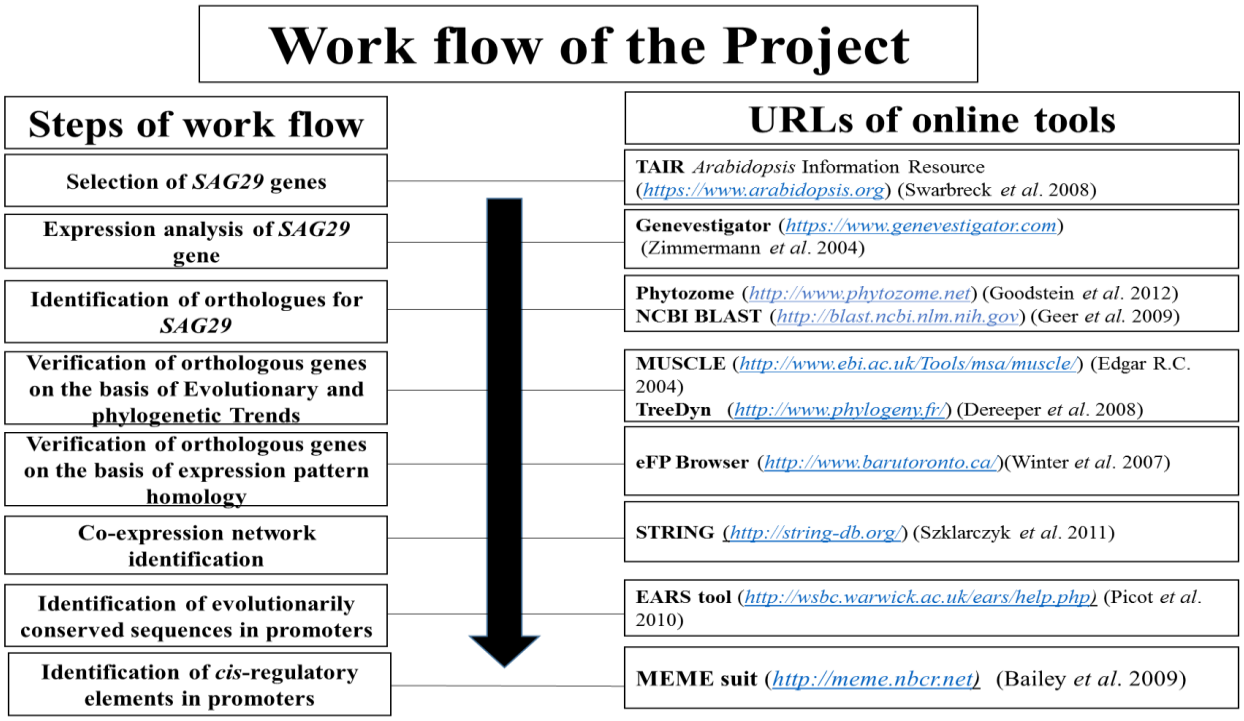


Figure 2.1 Workflow for *de-novo* Identification/Prediction of putative *cis*-regulatory elements (pCREs) in the 1kb promoter sequences of *SAG29* and their orthologues in selected plant species

2.2 Wet lab work

2.2.1 Testing early growth response upon selenium application in plants

Seeds of soybean (*Glycine max*) were soaked in water for 3 hours to prevent dormancy then seeds are grown in sand pots. Seeds of soybean plants were grown for maximum ten days though giving the proper amount of water daily to avoid the seeds from drying out. Three biological replicates are used with control for each experiment, and were provided with proper amount of water, and remaining three biological replicates were treated on eleventh day of germination with 100µM sodium selenite (Na2SeO3). Various growth parameters were assessed during early developmental stages, in the set of three biological replicates, as described (Pauwels *et al.*, 2009), such as shoot length and total chlorophyll content and auxin concentration. Experimental data of three biological replicates were statistically analyzed using ANOVA.

2.2.1.1 Shoot length

Soybean shoot length of three biological replicates of both control and treated were measured with help of scale on seventh day after treatment. Four plants were taken from each biological replicate. Next the data was analyzed.

2.2.1.2 Chlorophyll content

Chlorophyll content of soybean plants were determined with chlorophyll meter on seventh day after treatment. Three biological replicates were made for control and treated plants. Two leaves were taken from each biological replicates of control and treated plants.

2.2.1.3 Exogenous auxin (IAA) level

Materials required:

Samples were centrifuged first at 10,000 rpm for 2 minutes and then salkoski reagent was added to each Sample, kept in dark period for 30 minutes and observed on photo -

spectrometer on OD540nm.

Table 2.1 Shows that the Salkoski reagent constituents.

Salkoski reagent constituents	Concentration
Perchloric acid	50 % conc with 50 % H2O
Fecl3	0.8125 g/10 mL of H20

Table 2.2 shows the constituent and quantity for the determination of IAA content.

Constituent	Volume
Salkoski reagent	2 ml
Sample	1 ml

Procedure:

O.3 gram sample was taken and grind it with 4 ml of distal water. Then 1 ml of sample was taken in appendorf.Sample was centrifuged at 10,000 rpm for 2 minutes.Take the supernatant in test tube.2 ml salkoski reagent was added to the test tube with the help of dropper.

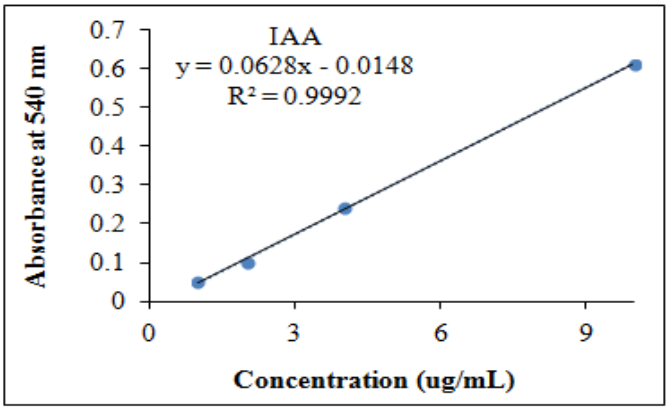


Figure 2.2 Shows the standard curve Auxin (IAA)

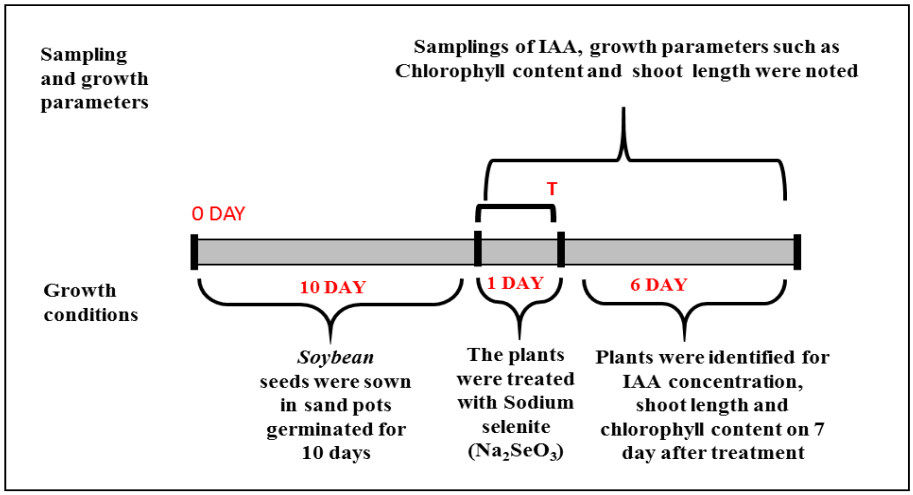


Figure 2.3 Work plan for Samplings of IAA, growth parameters such as Chlorophyll content and shoot length

3: Results and Discussion

3.1.1 Selection and Expression analysis of the *SAG 29*

The expression analysis of the members of *SAG 29* by a well-recognized search engine was used called as Genevestigator. It is a high quality bioinformatics search database for investigating gene expression. Gene *SAG29* were analyzed at anatomical and developmental level by recovering the standards of expression from affymetrix array database from Genevestigator. For this, microarray expression data was obtained using *Arabidopsis* Gene Chip platform. For *SAG29* and its selected orthologues gene identifiers remained used as query sequences and carry out explorations in the Gene Chip platform of Genevestigator, where microarray data was analyzed only for the wild type background. The data could be retrieved for *Arabidopsis thaliana (SAG29; AT5G13170), Glycine max (Glyma.04G198500)* and *Zea mays (GRMZM5G872392_T01)*. The data obtained in various plants developmental stages was saved as graphs as shown in (**Fig. 3.1**, **3.2** and **3.3**). In developmental stage specific analysis, throughout life cycle of all three genes presented modest expression. The gene expression of *SAG29* was greatly

upregulated at the early growth stages (from development to 2-leaf phase) in situation of *Glycine max* shadowed by modest expression through the entire life phases, and again higher during dough stage. High expression in the initial as well as finally dough stages in *Glycine max* conceivably provides higher requirement of N-remobilization from source to sink translations, at these phases of lifespan. Generally, the exploration specifies that the gene must perform some vital parts in retentive welfare conditional stages of the *Glycine max* during development of lifespan.

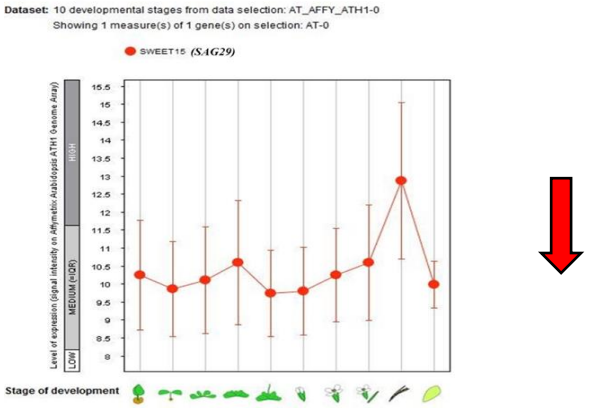


Figure 3.1 Developmental expression patterns of *SAG29* in *Arabidopsis thaliana*

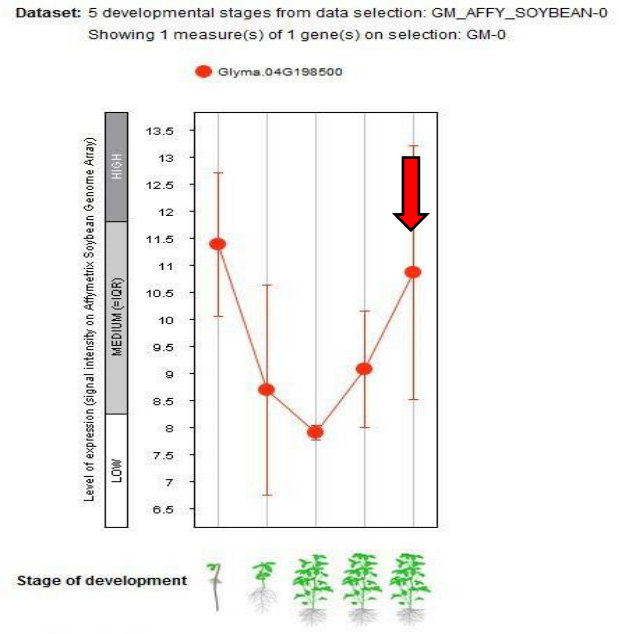


Figure 3.2 Developmental expression patterns of *SAG29 orthologue* gene (*Glyma.04G198500*) in *Glycine max*

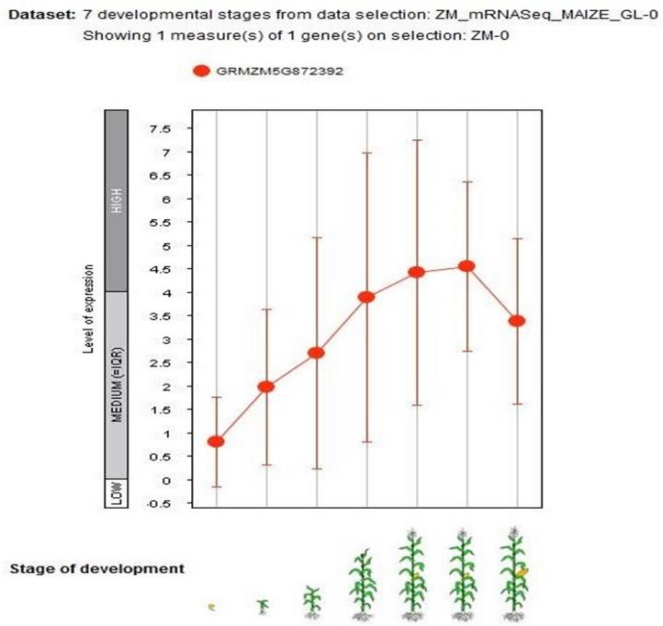


Figure 3.3 Developmental expression patterns of *SAG29* orthologue (*GRMZM5G872392)* in *Zea mays*

3.1..2 Genome-wide Identification for Orthologues of *SAG29 (AT5G13170)*

As a result of the search for orthologues of selected gene; *AT5G13170*, using nucleotide sequence of the selected gene i.e. *AT5G13170* as query. For the reference gene *AT5G13170* Ninteen orthologues were selected from Amino acid Identity 79.5% to 55.8% as shown in (**Table 3.1**). Out of these 19 orthologous from different plants species two orthologues namely *Glycine max* and *Zea mays* were selected for the reference genes due to their high valued economic importance and availability of their specialized online tools and databases for elucidating gene regulatory network.

Table 3.1 Representing *AT5G13170* orthologous genes with GeneInfo Identifiers from different plant species Data obtained from NCBI Blast search and Phytozome

Gene Info Identifier	Source Organism	AA Identity	CDs Length	AA Length
AT5G13170	*Arabidopsis thaliana*	100	879	292
Gorai.005G195100	*Gossypium raimondii*	79.5%	867	288

Thecc1EG008493	*Theobroma cacao*	77.4%	864	287
Ciclev10032218m.g	*Citrus clementina*	73.3%	918	305
Manes.13G006800	*Manihot esculenta*	71.9%	858	285
Manes.12G006700	*Manihot esculenta*	70.5%	267	288
evm.TU.supercontig_99.6	*Carica papaya*	69.2%	237	278
Potri.001G060900	*Populus trichocarpa*	67.5%	849	282
29929.m004599	*Ricinus communis*	67.1%	819	272
Glyma.04G198500	*Glycine max*	66.4%	849	282
Prupe.1G220700	*Prunus persica*	65.1%	834	277
GSVIVG01000938001	*Vitis vinifera*	64.4%	870	289
Brast04G189800	*Brachypodium stacei*	64.4%	921	306
Pavir.Hb01479	*Panicum virgatum*	64.4%	927	308
SapurV1A.0173s0030	*Salix purpurea*	63.4%	810	269
Aqcoe3G159400	*Aquilegia coerulea*	63.0%	903	300
Migut.D01285	*Mimulus guttatus*	62.3%	831	276
Sevir.8G137700	*Setaria viridis*	62.3%	891	296
GSMUA_Achr10G11880_001	*Musa acuminata*	61.6%	1005	334
GRMZM5G872392_T01	*Zea mays*	55.8%	1002	333

3.1.3 Verification of Orthologous genes on the basis of Evolutionary and phylogenetic Trends (Construction of Evolutionary Trees)

The selected closest members of orthologous genes of the reference gene from different species were selected to extract the putative 1000-bp promoter region counted from transcription initiation codon (ATG) and CDSs. The phylogenetic trees were constructed by using MUSCLE multiple sequence alignment tool then the phylograms were screened by TreeDyn. The Phylogram assisted in identification and confirmation of a number of orthologous genes which were supported with high bootstrap values. Maximum likelihood bootstrap consensus phylogeny of CDS and 1kb Promoter sequences of *SAG29* and its orthologous genes was carried out on Neighbor-joining distance tree based parameter with bootstrap support and node length is given along the branches which also verified the closest orthologous members of Reference gene as shown in the following (**Fig 3.4, 3.5, 3.6** and **3.7**). Although, the importance of *SAG29* gene is elucidated earlier by researchers, however the evolutionary events and conserved regulatory sequences in promoter region remained unclear so far. It is well known that the promoter sequences and CDS of the orthologous genes are expected to show close association among various plant species in order to have conserved developmental expression as well as biological role. High bootstrap values supported by the phylogram helped in prediction of consensus promoter sequence among and within species indicating conservativity at expression level. A phylogenetic tree thus constructed with high confidence using neighbor-joining method, exemplifies their ancestral relationships. Tree is built aligned each sequence using the reference gene 1-kb promoter and CDSs from *Arabidopsis thaliana* as query and their most similar sequences determined by BLAST.

3.1.4 Screening of the reference gene and its Orthologues on the basis of Expression Pattern Homolgy by eFP Browser

The selected orthologues *Glycine max* (*Glyma.04G198500*) and *Zea mays* (*GRMZM5G872392_T01*) along with their reference gene were assessed on the basis of expression

pattern homology at different developmental stages and threshold by using Bio-Analytic Resource for Plant Biology/eFP Browser as shown in the following (**Fig. 3.8, 3.9** and **3.10**). The reference gene *AT5G13170* and its orthologues shows moderate expression at development stage as shown in **(Table 3.2).**

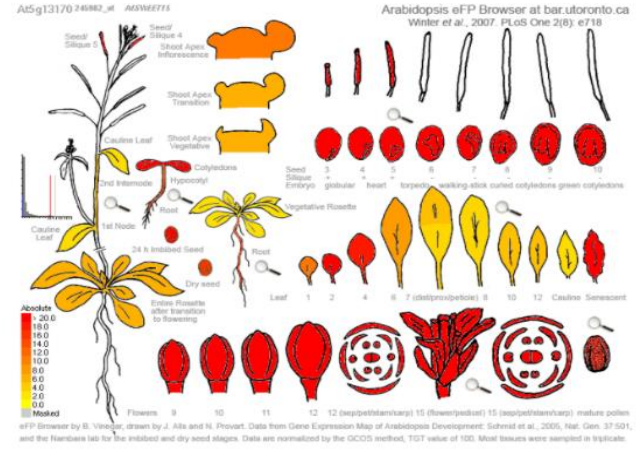


Figure 3.8 Screening of the reference gene *SAG29* on the basis of tissue based differential Expression Pattern Homology

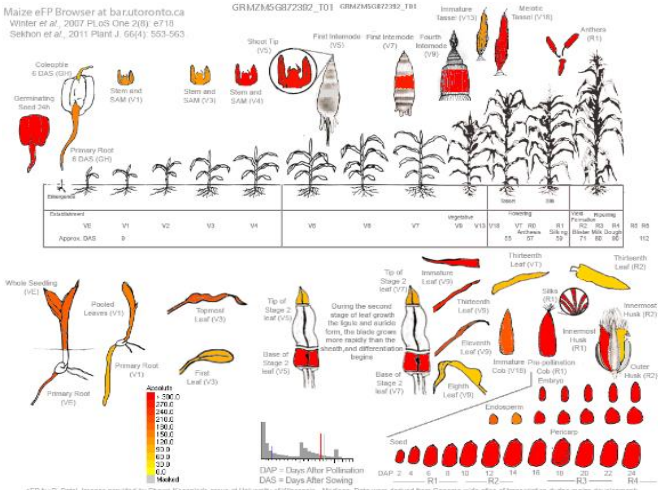


Figure 3.9 Screening of *SAG29* orthologue (GRMZM5G872392_T01) Zea mays on the basis of tissue based differential Expression Pattern Homology

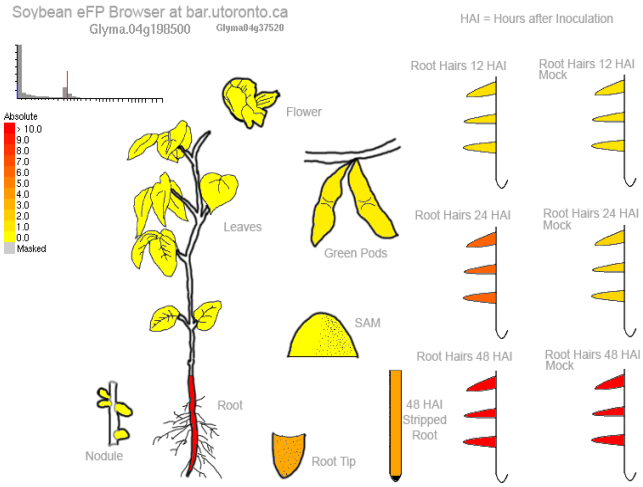


Figure 3.9 Screening of *SAG29* orthologue *Glycine max* (*Glyma.04G198500*) on the basis of tissue based differential Expression Pattern Homology

Table 3.2 Screening of the reference gene *AT5G13170* orthologous genes on the basis of Expression Pattern Homology

Expressional similarity	***Arabidopsis thaliana* (*AT5G13170*)**	***Zea mays* (*GRMZM5G872392_T01*)**	***Glycine max* (*Glyma.04G198500*)**
High Expression	Flower, Roots, Seeds	Seeds and leaves	Roots
Low Expression	Stem and leaves	Stem and SAM	Stem and leaves

3.1.7 Identification of the known *cis*-regulatory elements in 1000 bp sequences of Reference genes and its orthologues in promotor and introns

For the exploration of the identified *cis*- regulatory elements in 1kb promoter and introns sequence region of the reference gene and its selected orthologues *Glycine max* (*Glyma.04G198500*) and *Zea mays* (*GRMZM5G872392_T01*) a database named PlantCARE was accessed. Different regulatory elements which were represented by consensus sequences, positional matrices and individual sites on

individual promoter and introns sequences was retrieved from the database. The list of these common *cis*-regulatory elements are given in the following (**tables 3.3** and **3.4**) for promotor and introns respectively.

Table 3.3 the list of common known *cis*-regulatory elements in 1000 bp Sequences of Reference genes *SAG29* (*At5g13170*) and its Orthologues for promotor.

Motif Name	Sequence	Site Name	Organism	Position	Gene
cis-acting element involved in the abs*cis*ic acid responsiveness	TACGTG	ABRE	*Arabidopsis thaliana*	109 and 349	AT5G13170
				136 and 163	Glyma.04G198500
				194 and 409	GRMZM5G872392_T01
cis-acting regulatory element involved in light responsiveness	CACGTA	G-Box	*Pisum sativum*	758	AT5G13170
				136	Glyma.04G198500
				277	GRMZM5G872392_T01
core promoter element around -30 of transcription start	TAATA	TATA-box	*Oryza sativa*	162	AT5G13170
				541	Glyma.04G198500
				618	GRMZM5G872392_T01
Light responsive element	CC(G/A)CC C	Sp1	*Zea mays*	191	AT5G13170
				512	Glyma.04G198500
				574 and 237	GRMZM5G872392_T01
cis-acting regulatory element required for endosperm expression	GTCAT	SKN-1 Motif	*Oryza sativa*	208 and 691	AT5G13170
				406	Glyma.04G198500
				185,1011 and 355	GRMZM5G872392_T01

Table 3.4 the list of common known *cis*-regulatory elements in 1000 bp Sequences of Reference genes *At5g13170* and its Orthologues for intron.

Motif Name	Sequence	Site Name	Organism	Position	Gene
cis-acting element involved in the abs*cis*ic acid responsiveness	TACGTG	ABRE	*Arabidopsis thaliana*	109 and 349	AT5G13170
				136 and 163	Glyma.04G198500
				194 and 409	GRMZM5G872392_T01
cis-acting regulatory element involved in light responsiveness	CACGTA	G-Box	*Pisum sativum*	758	AT5G13170
				136	Glyma.04G198500
				277	GRMZM5G872392_T01
core promoter element around -30 of transcription start	TAATA	TATA-box	*Oryza sativa*	162	AT5G13170
				541	Glyma.04G198500
				618	GRMZM5G872392_T01
Light responsive element	CC(G/A)CCC	Sp1	*Zea mays*	191	AT5G13170
				512	Glyma.04G198500
				574 and 237	GRMZM5G872392_T01
cis-acting regulatory element required for endosperm expression	GTCAT	SKN-1 Motif	*Oryza sativa*	208 and 691	AT5G13170
				406	Glyma.04G198500
				185,1011 and 355	GRMZM5G872392_T01

3.1.6 Identification of Evolutionarily Conserved Sequences within the Promoter of Reference genes and its Orthologues by EARS tool

3.1.6.1 Evolutionarily conserved sequences within the promoter

As promoter regions are estimated to be conserved throughout evolution and are functionally important for gene expression and, a newly developed comparative genomics technique were applied to identify regulatory elements that are conserved between *SAG29* orthologues from reserved species. The region containing the conserved motif reveals a significant level of conservation among *Arabidopsis* and other species enlisted in (**Table 3.5**), *Glycine max* (*Glyma.04G198500*) and *Zea mays* (*GRMZM5G872392_T01*) are shown in (**Fig. 3.10, 3.11** and **3.12**). The region of significant conservation consensus sequence (CNS) (indicated with curly black colored bracket) .Consensus regions may possess conserved motifs functionally important for gene expression as novel *cis*-regulatory elements. This finding has potential to be proved by experimental approaches.

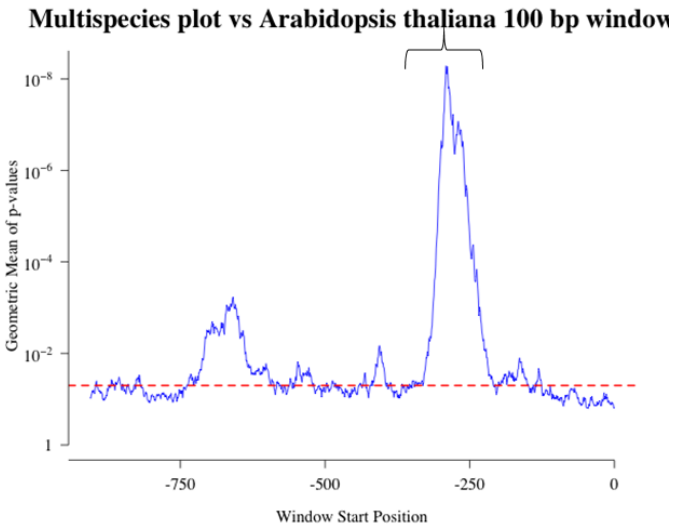


Figure 3.10 showing the multispecies plot result of EAR tool

for the Identification of Evolutionarily Conserved Sequences within the Promoter of *SAG29* (*AT5G13170)* and its orthologous genes. The significance threshold of P = 0.05 are used showing by spotted red color lines. And the window has a highly conserved match in the other species are presenting by peaks above this threshold. (A) Consensus sequence (CNS) is highlighted with BLACK curly bracket at the location of conservativity. Due to their low complexity, consensus has given a high conservation score.

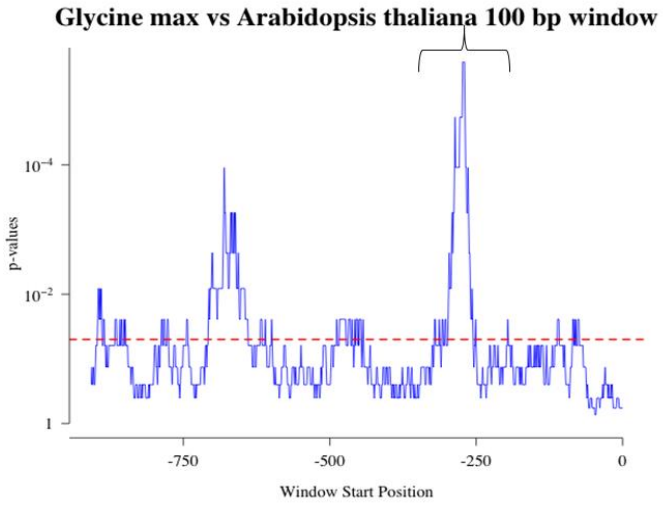


Figure 4.11 Showing the result of EAR tool for the Identification of Evolutionarily Conserved Sequences within the Promoter of *SAG29 (AT5G13170)* and its orthologous gene *Glyma.04G198500 (Glycine max)*. The significance threshold of P = 0.05 are used showing by spotted red color lines. And the window has a highly conserved match in the other species are presenting by Peaks above this threshold .Consensus sequence (CNS) is highlighted with BLACK curly bracket.

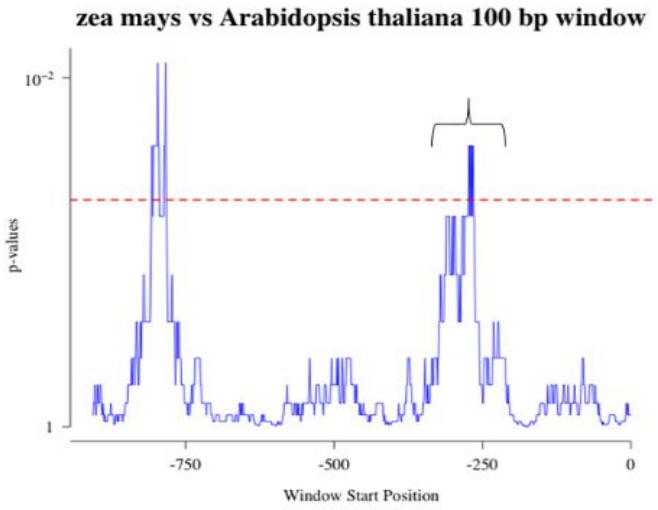


Figure 4.12 showing the multispecies plot result of EAR tool for the Identification of Evolutionarily Conserved Sequences within the Promoter of *SAG29* (*AT5G13170)* and its orthologous gene. The significance threshold of P = 0.05 are used showing by spotted red color lines. And the window has a highly conserved match in the other species are presenting by peaks above this threshold. (A) Consensus sequence (CNS) is highlighted with BLACK curly bracket at the location of conservativity. Due to their low complexity, consensus has given a high conservation score

Table 3.5 Showing the position of conserved motif start site of the reference genes and their orthologues for promotor.

Gene Info Identifier	Source Organism	Motif location
AT5G13170	*Arabidopsis* thaliana	275-291
Glyma.04G198500	*Glycine max*	271-291
GRMZM5G872392_T01	*Zea mays*	252-272

3.1.6.2 Evolutionarily conserved sequences within the introns

Regulatory regions are also occur in the introns of genes and are not restricted to the promoter. To find regulatory elements that are conserved between *SAG29* orthologues from reserved species we also performed comparative genomics technique. The region comprising the conserved motifs exhibits a significant level of conservation between *Arabidopsis* and distant species enlisted in (**Table 3.6**), including *Glycine max* (*Glyma.04G198500*) and several others as shown in (**Fig. 3.13** to **2.20**). The region of significant conservation is consensus (indicated with curly black colored bracket), consensus regions may possess conserved motifs functionally important

for gene expression as novel *cis*-regulatory elements. This finding has potential to be proved by experimental approaches.

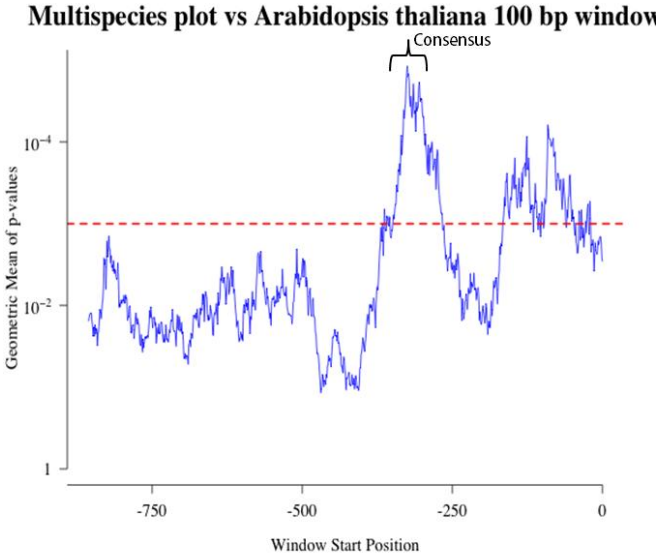


Figure 3.13 showing the multispecies introns of *SAG29* (*AT5G13170)* and its orthologous gene. The significance threshold of P = 0.001 are used showing by spotted red color lines. And the window has a highly conserved match in the other species are presenting by peaks above this threshold. (A) Consensus sequence (CNS) is highlighted with BLACK curly bracket.

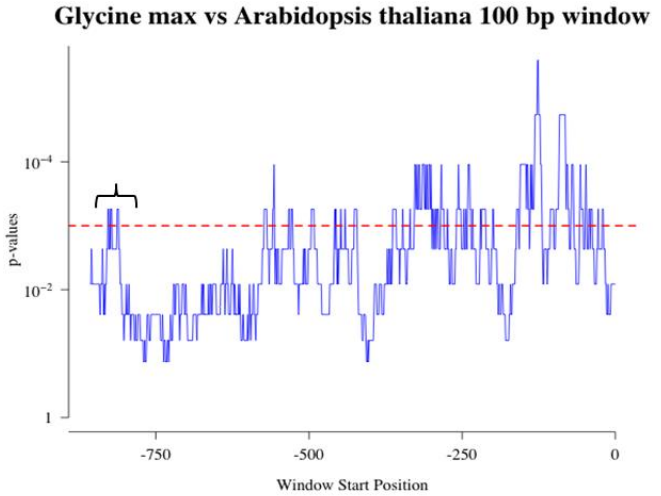


Figure 3.14 Showing the result of EAR tool for the Identification of Evolutionarily Conserved Sequences within the Introns of *SAG29 (AT5G13170)* and its orthologous gene *Glyma.04G198500 (Glycine max)* The significance threshold of P = 0.001 are used showing by spotted red color lines. And the window has a highly conserved match in the other species are presenting by Peaks above this threshold .Consensus sequence (CNS) is highlighted with BLACK curly bracket.

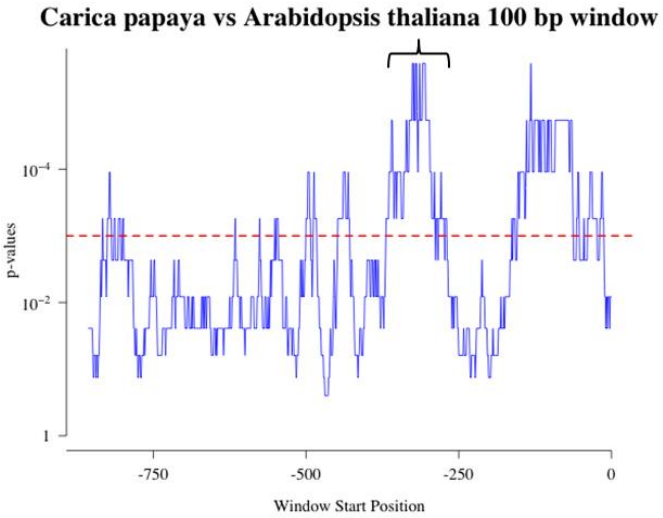


Figure 3.15 Showing the result of EAR tool for the Identification of Evolutionarily Conserved Sequences within the Introns of *SAG29 (AT5G13170)* and its orthologous gene *evm.TU.supercontig_99.6* (*Carica papaya)*. The significance threshold of P = 0.001 are used showing by spotted red color lines. And the window has a highly conserved match in the other species are presenting by Peaks above this threshold .Consensus sequence (CNS) is highlighted with BLACK curly bracket.

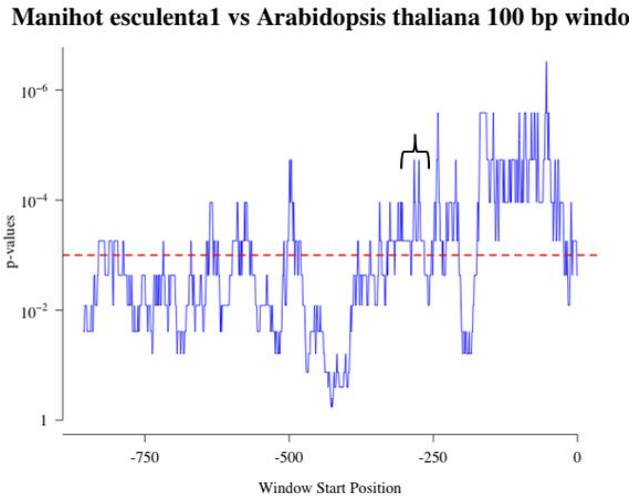


Figure 3.16 Showing the result of EAR tool for the Identification of Evolutionarily Conserved Sequences within the Introns of *SAG29* (*AT5G13170)* and its orthologous gene *Manes.13G006800* (*Manihot esculenta)*. The significance threshold of P = 0.001 are used showing by spotted red color lines. And the window has a highly conserved match in the other species are presenting by Peaks above this threshold .Consensus sequence (CNS) is highlighted with BLACK curly bracket.

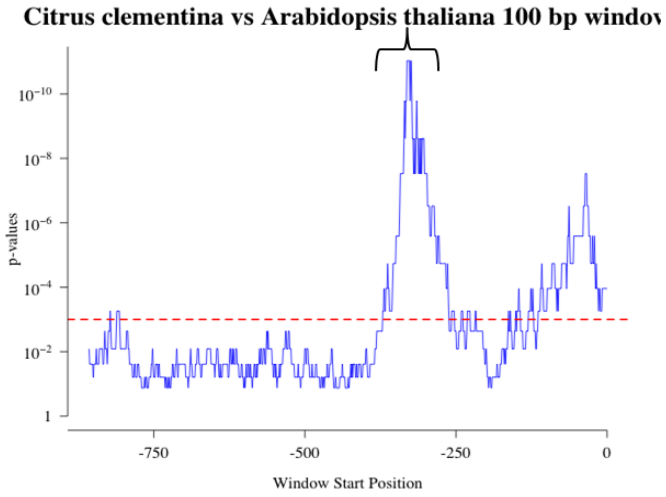


Figure 3.17 Showing the result of EAR tool for the Identification of Evolutionarily Conserved Sequences within the Introns of *SAG29* (*AT5G13170)* and its orthologous gene *Ciclev10032218m.g* (*Citrus clementina).* The significance threshold of P = 0.001 are used showing by spotted red color lines. And the window has a highly conserved match in the other species are presenting by Peaks above this threshold .Consensus sequence (CNS) is highlighted with BLACK curly bracket.

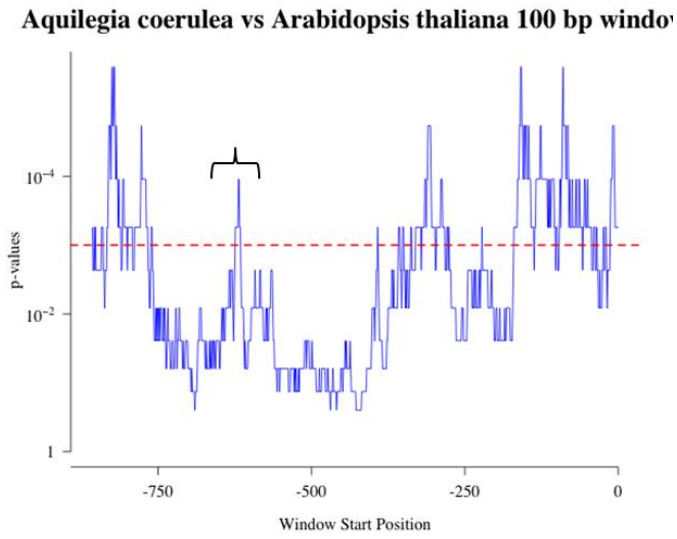


Figure 3.18 Showing the result of EAR tool for the Identification of Evolutionarily Conserved Sequences within the Introns of *SAG29* (*AT5G13170)* and its orthologous gene *Aqcoe3G159400* (*Aquilegia coerulea).* The significance threshold of P = 0.001 are used showing by spotted red color lines. And the window has a highly conserved match in the other species are presenting by Peaks above this threshold .Consensus sequence (CNS) is highlighted with BLACK curly bracket

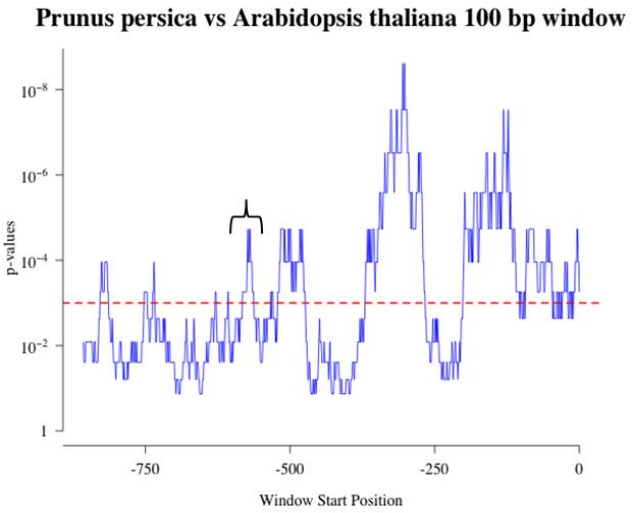


Figure 3.19 Showing the result of EAR tool for the Identification of Evolutionarily Conserved Sequences within the Introns of *SAG29* (*AT5G13170)* and its orthologous gene *Prupe.1G220700* (*Prunus persica).* The significance threshold of P = 0.001 are used showing by spotted red color lines. And the window has a highly conserved match in the other species are presenting by Peaks above this threshold .Consensus sequence (CNS) is highlighted with BLACK curly bracket.

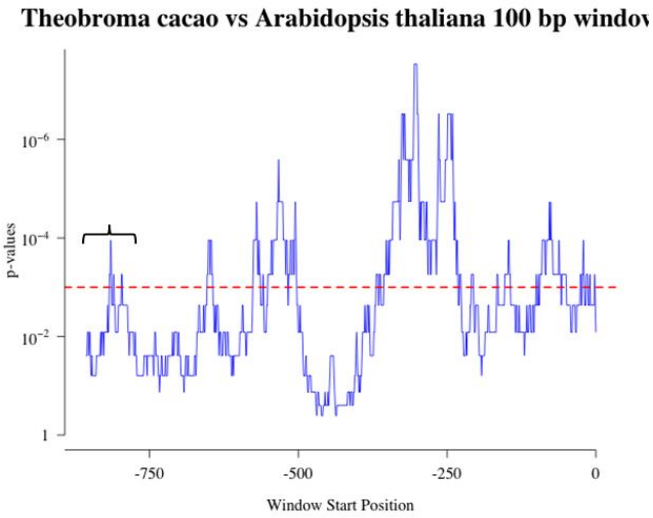


Figure 3.20 Showing the result of EAR tool for the Identification of Evolutionarily Conserved Sequences within

the Introns of *SAG29* (*AT5G13170)* and its orthologous gene Thecc1EG008493 (*Theobroma cacao).* The significance threshold of P = 0.001 are used showing by spotted red color lines. And the window has a highly conserved match in the other species are presenting by Peaks above this threshold .Consensus sequence (CNS) is highlighted with BLACK curly bracket.

Table 3.6 showing the position of conserved motifs start site of the reference genes and their orthologues for introns

Gene Info Identifier	Source Organism	Motif location
AT5G13170	*Arabidopsis* thaliana	251-258
Glyma.04G198500	*Glycine max*	811-818
evm.TU.supercontig_99.6	*Carica papaya*	334-341
Manes.13G006800	*Manihot esculenta*	347-354
Ciclev10032218m.g	*Citrus clementina*	400-407
Aqcoe3G159400	*Aquilegia coerulea*	611-618
Prupe.1G220700	*Prunus persica*	586-593
Thecc1EG008493	*Theobroma cacao*	786-793

3.1.7 Over-representative analysis for putative *cis*-elements in conserved non-Coding regions

3.1.7.1 Over-representative analysis for putative *cis*-elements in conserved non-coding regions of promoters

As we found two peaks above the threshold significance threshold of P = 0.05 which indicates that the window has a highly conserved match in the other species. Therefore to identify the conserved motif sequence at this region of promoters, MEME suit was used which revealed three consensus sequences between positions of −200 to -1000 of promoter (**Fig. 3.21**). Consensus sequence revealed a conserved motif in *SAG29* promoter form *Arabidopsis* thaliana between positions -275 and 295. The consensus sequence revealed the oligo sequence of *N*AA(T/A)[ATA]*NN*(C/G)*NN*G*N*AA(T/A)AA consisted of a ATA core sequence. The positional profiles of these motifs in orthologous species are shown in (**table 3.5**)

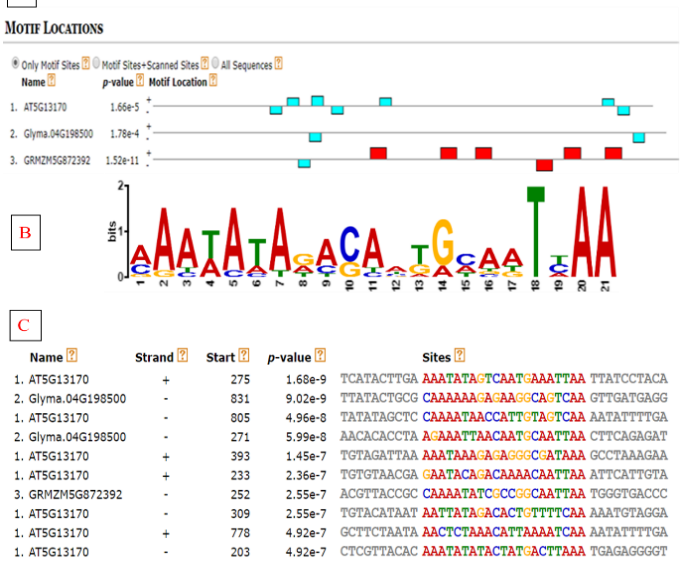


Figure 3.21 Cumulative conservation profile of the Reference gene *SAG29* (*AT5G13170)* promoter and orthologous promoters from distant species. (A) 1kb upstream of the translational start site of the *SAG29* gene were affiliated by the EARs tool for identification of the evolutionary conserved regulator regions. (B) Showing positional profile and consensus sequence with Logo within *SAG29* promoter from *Glycine max (Glyma.04G198500).*

3.1.7.2 Over-representative analysis for putative *cis*-elements in conserved non-coding regions of introns

As we found one peak above the threshold significance threshold of P = 0.001 which indicates that the window has a conserved match in the other species. Therefore to identify the conserved motif sequence at this region of introns, MEME suit was used which revealed one consensus sequences between positions of −200 to -1000 of introns (**Fig. 3.22**). Consensus sequence revealed a conserved motif in *SAG29* introns form *Arabidopsis* thaliana between positions -251 and -258. The consensus sequence revealed the oligo sequence of *N*[AGGT](G/A)AC consisted of a AGGT core sequence. The positional profiles of these motifs in orthologous species are shown in (**table 3.6**).

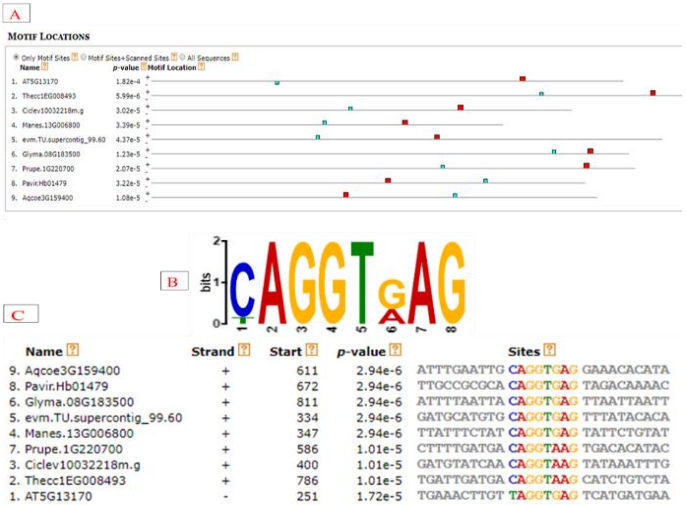


Figure 3.22 Cumulative conservation profile of the Reference gene *SAG29* (*AT5G13170)* introns and orthologous introns from distant species. (A) 1kb upstream of the translational

start site of the *SAG29* gene were affiliated by the EARs tool for identification of the evolutionary conserved regulator regions (B) Showing positional profile and consensus sequence with Logo within *SAG29* introns from various plant species.

3.1.8 Co-Expression Network Identification of Reference and Orthologous genes

An online database resource and a well-recognized "STRING" (search tool for the recovery of Interacting proteins of genes) (http://string-db.org/) was accessed to reveal the functional interactions of the proteins of reference gene and its orthologous. It provided a comprehensive data for both the predicted interactions as well as experimental information. The interacting proteins of reference gene and its selected orthologues are shown in the (**Fig 3.23, 3.24** and **3.25**) the lists of co-expressed proteins are given in the (**table 3.7**) as follows.

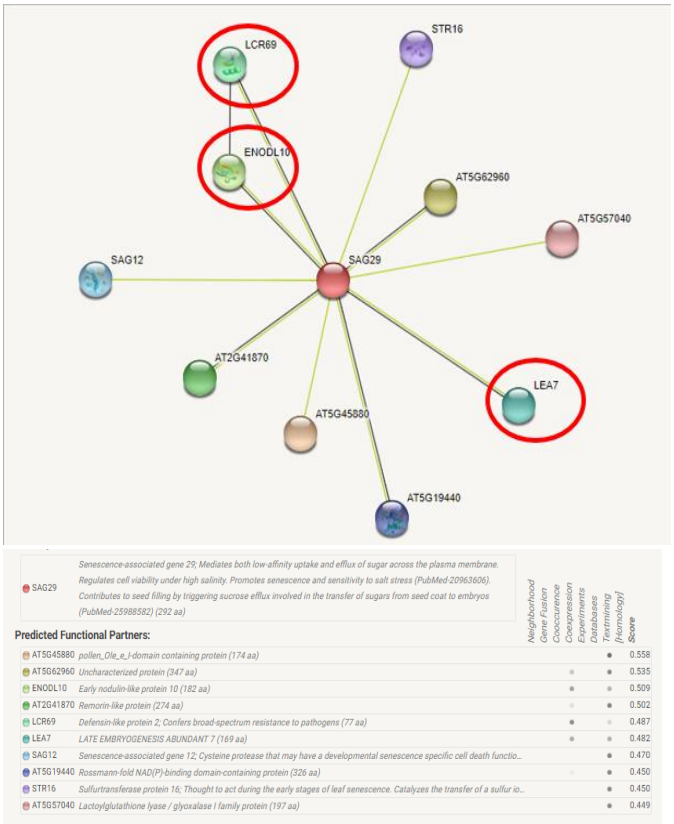
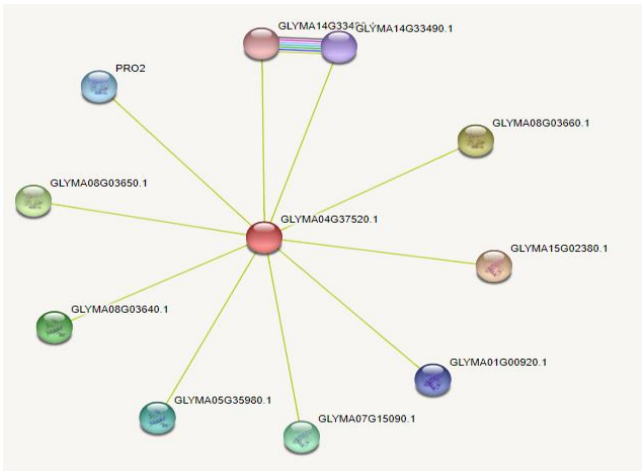


Figure 3.23 Co-Expression Network Identification of the reference gene *SAG29* (*AT5G13170)*, the co-expressed proteins encircled with red color.

.



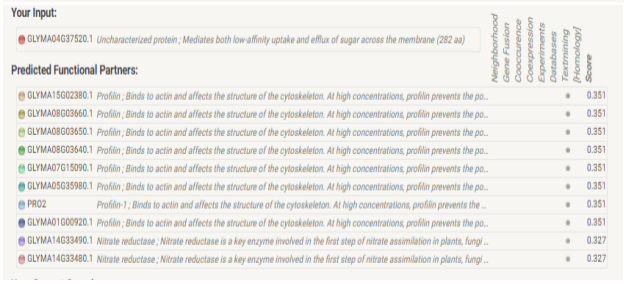
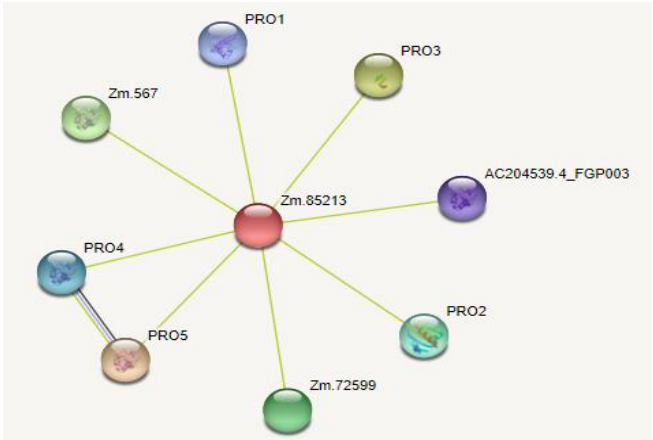


Figure 3.24 Co-Expression Network Identification of orthologue of the reference gene *SAG29, Glyma.04G198500* (*Glycine max)* showing all the uncharacterized co-expressed proteins.



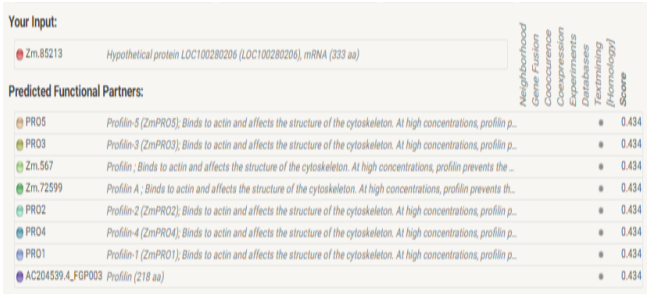


Figure 3.25 Co-Expression Network Identification of orthologue of the reference gene *SAG29*, *GRMZM5G872392_T01* (*Zea mays)* showing all the uncharacterized co-expressed proteins

Table 3.7 Co-Expression Network Identification of *AT5G13170* and its orthologous genes

Gene Name	Co expressed gene/Protein	Function
At5g13170 (Arabidopsis thaliana)	*ENODL10*	electron carrier activity
	LCR69	resistance response , resistance response to fungus, killing of cells of other organism
	LEA7	protein stabilization, response to freezing, response to water deprivation
Glyma.04G198500 (Glycine max)	Uncharacterized Protein	sugar transmembrane transporter activity
GRMZM5G872392_T01 (Zea mays)	Uncharacterized Protein	sugar transporter

3.2 Wet lab work:

3.2.1 Testing early growth response upon selenium application in Soybean:

3.2.1.1 Shoot length

The plants were treated with 4ml of Sodium selenite (100µM) on eleven day of germination and the shoot length were measured on seventh day after treatment by using scale. The obtained results show a maximal shoot length in control biological replicates as compared to the plants treated with sodium selenite (100µM). It was observed that the shoots length were reduced in plants growing in the sodium selenite (100µM) as show in (**Fig. 3.26**). Shoot and leaf have capacity to assimilate more selenium as compared to roots. (Zayed *et al.* 1998).

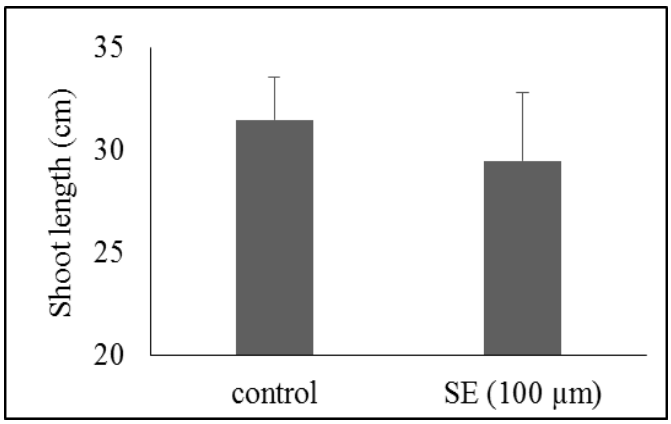


Figure 3.26 The impact of Sodium selenite on shoot length of Soybean (the amount of 4ml Sodium selenite (100µM) were given to the soybean plant and the shoot length were measured on seventh day after treatment).Bars represent the means ± Sodium selenite from 3 replicate experiments.

3.2.1.2 Chlorophyll content:

Chlorophyll content was determined with chlorophyll meter. Result indicate that soybean plants treated with 100 uM concentration of sodium selenite showed significant reduction in overall chlorophyll contents in comparison of control as shown in (**Fig. 3.27**). The development of leaf senescence is typically estimated by diverse physical signs such as fluctuations in chlorophyll contents or paling of leaf. The yield of soybean can be increased with application of selenium in soybean by avoiding chlorophyll reduction and sustaining elongated leaf capacity period (Djanaguiraman *et al.* 2004).

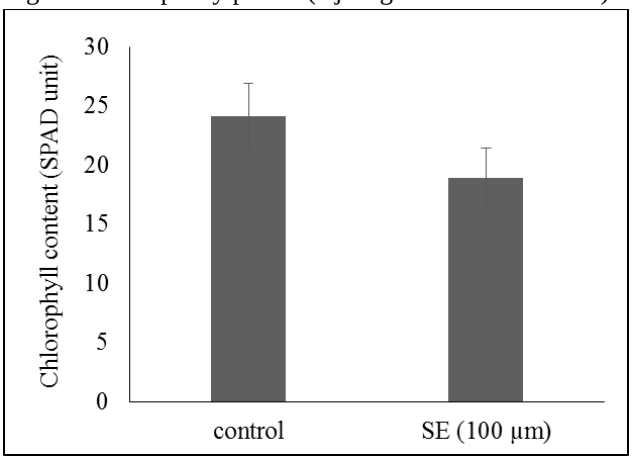


Figure 3.27 The impact of Sodium selenite on chlorophyll content of Soybean (the amount of 4ml Sodium selenite (100µM) were given to the soybean plant and the chlorophyll content was measure on seventh day after treatment). Bars

represent the means ± Sodium selenite from 3 replicate experiments.

3.2.1.3 Endogenouse auxin (IAA) level

Result indicate that 4 ml sodium selenite of 100µM concentration were used to given stress to soybean plants revealed maximum reduction in whole auxin (IAA) level as compared to their control as shown in (**Fig.3.28**).Calculating Auxin amount in each replicates by a procedure as mentioned earlier in materials and methods. And the protocol were used as shown in (**Table 2.1 and 2.2)**.The standered curve for auxin (IAA) concentration were shown in (**Fig 2.2**).

Sodium selenite harmfulness has reclaimed as a display to classify gene *SAG29* that effect on sodium selenite deposition through Soybean. . Result showed that sodium selenite exerted an influence on the development of soybean plants of each treated biological replicate liable on Sodium selenite concentration and the forms of *SAG*s appearance alter in retort to divergent treatments or situations.

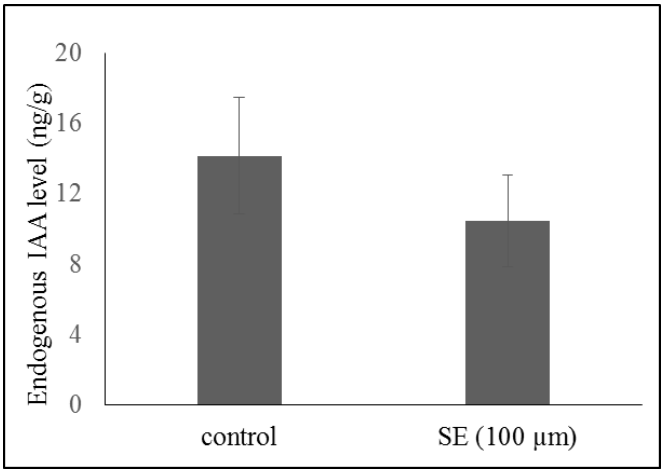


Figure 4.28 The impact of selenium salt (100µM) on Auxin (IAA) of Soybean (the amount of 4ml Sodium selenite (100µM) were given to the soybean plant and the IAA level was identified on seventh day after treatment). Bars represent the means ± Sodium selenite from 3 replicate experiments.

4: CONCLUSION

The computational analysis of evolutionarily conserved *cis* regulatory elements found in the *SAG29* gene of *Arabidopsis thaliana* and its orthologues namely *Glycine max* and *Zea mays* by phylogenetic approach has revealed highly conserved consensus regions with position specific motifs within -200 to -1000 bp of the promoter sequences, and a conserved consensus regions with position specific motifs within -200 to -1000 of the introns of *SAG29* and its orthologous distant plant species. To identify gene *SAG29,* sodium selenite toxicity's been used as a screen that impacts on sodium selenite accumulation by soybean. Result revealed the negative response of sodium selenite which induced senescence in early growth of soybean plants as shown in (**Fig 4.1**). However, in future one might consider to employ the yeast-one hybrid system to screen for upstream factors of transcription bound to the *SAG29* promoter, driving senescence activity and once the leaf senescence regulatory mechanisms are understood, it should be readily susceptive of devising more sophisticated leaf senescence ways to manipulation.

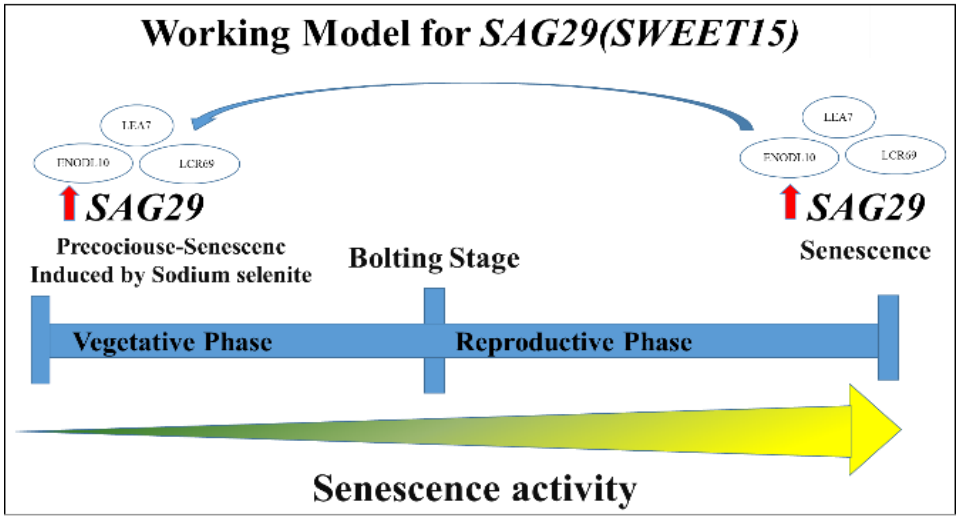


Figure 4.1 Model of negative effect of sodium selenite on soybean plants in response to senescence in early growth

REFERENCES

Agalou, A., Roussis, A. and Spaink, H.P., 2005. The *Arabidopsis* selenium-binding protein confers tolerance to toxic levels of selenium. *Functional Plant Biology*, *32*(10), pp.881-890.

Buchanan-Wollaston, V., Earl, S., Harrison, E., Mathas, E., Navabpour, S., Page, T. and Pink, D., 2003. The molecular analysis of leaf senescence–a genomics approach. *Plant biotechnology journal*, *1*(1), pp.3-22

Buchanan-Wollaston, V., Page, T., Harrison, E., Breeze, E., Lim, P.O., Nam, H.G., Lin, J.F., Wu, S.H., Swidzinski, J., Ishizaki, K. and Leaver, C.J., 2005. Comparative transcriptome analysis reveals significant differences in gene expression and signalling pathways between developmental and dark/starvation-induced senescence in *Arabidopsis*. *The Plant Journal*, *42*(4), pp.567-585

Balazadeh, S., Siddiqui, H., Allu, A.D., Matallana-Ramirez, L.P., Caldana, C., Mehrnia, M., Zanor, M.I., Köhler, B. and Mueller-Roeber, B., 2010. A gene regulatory network controlled by the NAC transcription factor ANAC092/AtNAC2/ORE1 during salt-promoted senescence. *The Plant Journal*, *62*(2), pp.250-264.

Breeze, E., Harrison, E., McHattie, S., Hughes, L., Hickman, R., Hill, C., Kiddle, S., Kim, Y.S., Penfold, C.A., Jenkins, D. and Zhang, C., 2011. High-resolution temporal profiling of transcripts during *Arabidopsis* leaf senescence reveals a distinct chronology of processes and regulation. *The Plant Cell*, *23*(3), pp.873-894.

Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W. and Noble, W.S., 2009. MEME SUITE: tools for motif discovery and searching. *Nucleic acids research*, p.gkp335.

Chen, L.Q., Hou, B.H., Lalonde, S., Takanaga, H., Hartung, M.L., Qu, X.Q., Guo, W.J., Kim, J.G., Underwood, W., Chaudhuri, B. and Chermak, D., 2010. Sugar transporters for intercellular exchange and nutrition of pathogens. *Nature*, *468*(7323), pp.527-532.

Chen, L.Q., Lin, I.W., Qu, X.Q., Sosso, D., McFarlane, H.E., Londoño, A., Samuels, A.L. and Frommer, W.B., 2015. A cascade of sequentially expressed sucrose transporters in the seed coat and endosperm provides nutrition for the *Arabidopsis* embryo. *The Plant Cell*, *27*(3), pp.607-619.

Edgar, R.C., 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC bioinformatics*, *5*(1), p.113.

Geer, L.Y., Marchler-Bauer, A., Geer, R.C., Han, L., He, J., He, S., Liu, C., Shi, W. and Bryant, S.H., 2009. The NCBI biosystems database. *Nucleic acids research*, *38*(suppl_1), pp.D492-D496.

Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N. and Rokhsar, D.S., 2012. Phytozome: a comparative platform for green plant genomics. *Nucleic acids research*, *40*(D1), pp.D1178-D1186.

Hörtensteiner, S., 2006. Chlorophyll degradation during senescence. *Annu. Rev. Plant Biol.*, *57*, pp.55-77

Lim, P.O., Kim, H.J., Nam, H.G., 2007. Leaf Senescence, in Annual Review of Plant Biology. Annual Reviews, Palo Alto, pp. 115-136

Lohman, K.N., Gan, S., John, M.C. and Amasino, R.M., 1994. Molecular analysis of natural leaf senescence in *Arabidopsis* thaliana. *Physiologia Plantarum*, *92*(2), pp.322-328.

Lescot, M., Déhais, P., Thijs, G., Marchal, K., Moreau, Y., Van de Peer, Y., Rouzé, P. and Rombauts, S., 2002. PlantCARE, a database of plant *cis*-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic acids research*, *30*(1), pp.325-327.

Meng Yuan, S. W., 2013. Rice MtN3/Saliva/*SWEET* Family Genes and Their Homologs in Cellular Organisms. *molecular plant*, p. 665–674.

Matallana-Ramirez, L.P., Rauf, M., Farage-Barhom, S., Dortay, H., Xue, G.P., Dröge-Laser, W., Lers, A., Balazadeh, S. and Mueller-Roeber, B., 2013. NAC transcription factor ORE1 and senescence-induced BIFUNCTIONAL NUCLEASE1 (BFN1) constitute a regulatory cascade in *Arabidopsis*. *Molecular plant*, *6*(5), pp.1438-1452.

Nam, H.G., 1997. The molecular genetic analysis of leaf senescence. *Current Opinion in Biotechnology*, *8*(2), pp.200-207.

Picot, E., Krusche, P., Tiskin, A., Carré, I. and Ott, S., 2010. Evolutionary analysis of regulatory sequences (EARS) in plants. *The Plant Journal*, *64*(1), pp.165-176.

Pauwels L, Inzé D and Goossens A. 2009.Jasmonate-inducible gene: what does it mean?. Trends in plant science, 14(2), pp.87-91.

Peña, D. G. A. S., 2015. "CARACTERIZACIÓN BIOINFORMÁTICA DE LA PROTEÍNA *SAG29* DE *Arabidopsis* thaliana INVOLUCRADA EN LA RESPUESTA AL ESTRÉS SALINO. *"CARACTERIZACIÓN BIOINFORMÁTICA DE LA PROTEÍNA SAG29 DE Arabidopsis thaliana INVOLUCRADA EN LA RESPUESTA AL ESTRÉS SALINO.*

Szklarczyk, D., Franceschini, A., Kuhn, M., Simonovic, M., Roth, A., Minguez, P., Doerks, T., Stark, M., Muller, J., Bork, P. and Jensen, L.J., 2011. The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. Nucleic acids research, 39(suppl 1), pp.D561-D568.

Seo, P.J., Park, J.M., Kang, S.K., Kim, S.G., and Park, C.M. (2011). An Arabidopsis senescence-associated protein SAG29 regulates cell viability under high salinity. Planta 233: 189–200.

Schauer, S.E., Schlu¨ter, P.M., Baskar, R., Gheyselinck, J., BolaO`os, A., Curtis,M.D. and Grossniklaus, U. (2009) Intronic regulatory elements determinethe divergent expression patterns of AGAMOUS-LIKE6 subfamily mem-bers in *Arabidopsis*. Plant J. 59, 987–1000.

Swarbreck, D., Wilks, C., Lamesch, P., Berardini, T.Z., Garcia-Hernandez, M., Foerster, H., Li, D., Meyer, T., Muller, R., Ploetz, L. and Radenbaugh, A., 2008. The *Arabidopsis* Information Resource (TAIR): gene structure and function annotation. *Nucleic acids research*, *36*(suppl 1), pp.D1009-D1014.

Weaver, L.M., Gan, S., Quirino, B. and Amasino, R.M., 1998. A comparison of the expression patterns of several senescence-associated genes in response to stress and hormone treatment. *Plant molecular biology*, *37*(3), pp.455-469.

Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G.V. and Provart, N.J., 2007. An "Electronic Fluorescent Pictograph" browser for exploring and analyzing large-scale biological data sets. PloS one, 2(8), p.e718.

Xu, F., Meng, T., Li, P., Yu, Y., Cui, Y., Wang, Y., Gong, Q. and Wang, N.N., 2011. A soybean dual-specificity kinase, GmSARK, and its *Arabidopsis* homolog, AtSARK, regulate leaf senescence through synergistic actions of auxin and ethylene. *Plant physiology*, *157*(4), pp.2131-2153.

Zhonghai Li, J. P. W. G., 2012. Gene Network Analysis and Functional Studies of Senescence-associated Genes Reveal Novel Regulators of *Arabidopsis* Leaf Senescence. *Gene Network*

Zimmermann, P., Hirsch-Hoffmann, M., Hennig, L. and Gruissem, W., 2004. GENEVESTIGATOR. *Arabidopsis* microarray database and analysis toolbox. *Plant physiology*, *136*(1), pp.2621-2632.

Zeng, F., Roslinsky, V. and Cheng, B., 2017. Mutations in the promoter, intron and CDS of two FAD2 generate multiple alleles modulating linoleic acid level in yellow mustard. *Scientific Reports*, *7*(1), p.8284