J. Plant Biochemistry & Biotechnology Vol. 18(1), 13-20, January 2009

In Silico and Deletion Analysis of Upstream Promoter Fragment of S-Adenosyl Homocysteine Hydrolase (*SAHH1)* Gene of *Arabidopsis* Leads to the Identification of a Fragment Capable of Driving Gene Expression in Developing Seeds and Anthers

T P Sujatha, C Sivanandan, S R Bhat and R Srinivasan*
National Research Center on Plant Biotechnology, Indian Agricultural Research Institute, New Delhi 110 012, India

S-adenosyl homocysteine hydrolase (SAHH) is a key enzyme in methylation metabolism of eukaryotes. A 1585 bp fragment upstream to ATG of *SAHH1* gene, was fused with a promoter-less β-Glucuronidase (*GUS*) gene and mobilized into *Arabidopsis* by *Agrobacterium*-mediated floral transformation to generate transgenic *Arabidopsis.* This fragment was found to drive constitutive expression of *GUS* in T_2 progeny of transgenic *Arabidopsis. In silico* analysis of the promoter region of *SAHH1* suggested the presence of several *cis*-regulatory motifs including seed-specific motifs as well as anther-specific motifs in the 376 bp (upstream to TSS of *SAHH1*) promoter fragment. Based on the partial deletion analysis carried out in the promoter region of *SAHH1* (At4g13940) this 376 bp promoter fragment was found to be capable of driving GUS expression in developing seeds and in some anthers/microspores.

***Key words*:** *SAHH1*, promoter, methylation, GUS, seed-specific.

S-adenosyl homocysteine hydrolase (SAHH), an enzyme involved in SAM-dependent methylation reactions, is widespread in eukaryotes (1). The expression of *SAHH* gene, though expected to be constitutive (2), was reported to increase in response to fungal elicitors, plant hormones, salinity, drought/dehydration and during developmental processes like lignification, fruit ripening and embryo/ endosperm development (3-10).

DNA methylation is necessary for homology-dependent gene silencing. The major function of SAHH enzyme in gene silencing in plants, is due to its role in methylation cycle. In *Arabidopsis,* two paralogues of *SAHH* gene *viz*., *SAHH1/HOG1* (homology-dependent gene silencing1), (At4G13940) and *SAHH2* (At3g13930) exist. Homozygous mutants of *AHH1*(EMBRYODEFECTIVE1395) are embryonic lethal in *Arabidopsis*. Unlike *SAHH1*, the homozygous mutant of *SAHH2* is not impaired in any aspect of growth, fertility and DNA methylation (11). A T-DNA insertion mutant of AtSAHH1 (*sahh1-1*) and the RNAi plants of AtSAHH2 (dsAtSAHH2) exhibited different degrees of phenotypic changes, including higher chlorophyll content, delayed flowering, senescence, and bushy architecture (12).

In our earlier study, a T-DNA tagged mutant M57 of *Arabidopsis,* with a promoter-less *GUS* gene, exhibiting intense GUS expression in the emerging lateral roots, the vascular region of roots and root tips of young seedlings, was characterized (13). A 452 bp promoter fragment lying immediately upstream of the *GUS* in T-DNA and in the intergenic region between two divergently transcribed genes namely *SAHH1* (At4G13940) and *SHMT4* (At4G13930), was isolated and cloned from *Arabidopsis thaliana* ecotype Columbia. The 452 bp promoter, which drives *GUS* expression in a pattern similar to that of mutant line 57 (14) is located about 1 Kb upstream to *SAHH1* gene, and the *GUS* gene driven by the promoter is in the same orientation as that of *SAHH1*. Two promoter fragments of 543 bp and 1181 bp upstream to the ATG of *SAHH1* were cloned separately, in promoter-less *GUS* vector, pBI101. The *GUS* expression driven by these promoters in transgenic *Arabidopsis* was found to be nearly constitutive, in all plant parts including the roots (13). This suggested that the root-specific expression driven by the 452 bp

*Corresponding author. E-mail : srinivasan53@gmail.com
Abbreviations: ABA-Abscisic acid; ADK-Adenosine kinase; *HOG1*-Homology-dependent gene silencing1; GUS-β-Glucuronidase; RNAi-RNA interference; SAHH-S-adenosyl homocysteine hydrolase; SAM-S-adenosyl methionine; TSS-transcription start site; UTR-untranslated region.

promoter fragment was probably independent of the downstream *SAHH1* gene. The promoter was therefore, designated as a cryptic root-specific promoter (AY601849). In the present study, a 1585 bp fragment (B1.5) upstream to ATG of *SAHH1* and another 376 bp (SAHD1) upstream to the experimental TSS of *SAHH1*, were cloned separately in pBI101 promoter-less vector upstream to the *GUS* gene, and the expression pattern of GUS was analysed in T_2 progeny.

Materials and Methods

Primer designing, PCR amplification and cloning of promoter fragments — Primers were designed manually using the sequence downloaded from TAIR (The *Arabidopsis* Information Resource) database available at *http://www.arabidopsis.org*. For directional cloning of 1.5 Kb fragment (B1.5), the restriction enzyme site for *Sal*I was added in the forward primer and that of *Bam*HI in the reverse primer. Restriction site for *Xba*I was added in the forward primer instead, for directional cloning of 376 bp SAHD1. At the extreme end, 2-3 base overhangs were incorporated for facilitating efficient restriction digestion within the recognition sequence. The sequences of the primers used in this study are given in Table 1. Primers R Prom 2 and AH Prom Bam were used to clone 1585 bp of B1.5, and primers F2 SAH and R1 SAH for cloning 376 bp of SAHD1.

CTAB method was followed for the isolation of genomic DNA from wild type *Arabidopsis* (15) and used as a template to amplify the promoter fragments. A 25 µl PCR reaction was set up with 1x PCR buffer (containing 1.5 mM $MgCl_2$), 200 µM of dNTP mix, 0.6 µM each of the corresponding forward and reverse primers, 1.5 U of *Taq* DNA polymerase, 150 ng of genomic DNA and sterile nuclease-free water to make up the volume. The promoter fragments were amplified at thermocycling conditions of 94°C for 3 min for initial denaturation, 94°C for 1 min for denaturation, 57°C for 45 sec for primer annealing, 72°C for 1 min for primer extension and 72°C for 10 min for final extension. Steps from denaturation to primer extension were repeated for 33 cycles. For amplifying 1.5 Kb fragment, primer extension was carried out at 68°C for 1.5 min, and PCR steps from denaturation to primer extension were repeated for 35 cycles.

Table 1. Sequences of primers used in the present study

Name	Sequence (5′- 3′)
R Prom 2	GTCGACGCTATGTATCACCCGGATG
AH Prom Bam	GGATCCGCTAGATCTGAGATTTG
F2 SAH	ACTCTAGAGTTAAACCAAGATTCAAC
R1 SAH	AGGGATCCTGTGCGAGGAGAAG
57RT-1	GTGCTCTAGATTGGGGTCCAGG
57RT-2	TCATCTTCCTCATGTGGTCGACC
ACT8F0	ATGAAGATTAAGGTCGTGGCA
ACT8R0	GACATCTCTCCAAACGCTGT

Note: *Sal*I (GTCGAC), *Xba*I (TCTAGA) and *Bam*HI (GGATCC) recognition sites are underlined.

The PCR products were cloned in TA cloning vector pGEM®-T Easy (Promega) and sequenced. The inserts from pGEM®-T Easy vector released after double digestion using *Sal*I/*Xba*I and *Bam*HI were purified from gel and cloned into the *Sal*I/*Xba*I and *Bam*HI sites upstream to the promoter-less reporter *GUS* gene in pBI101 vector. The recombinant clones were verified for the presence of insert by colony PCR as well as restriction digestion. The recombinant pBI101 carrying the deletion fragments were mobilized into *Agrobacterium tumefaciens* strain GV3101, and the presence of the promoter fragments in transformed colonies verified by colony PCR.

Plant material and growth conditions — Seeds of *A. thaliana* Ecotype Columbia were germinated on Murashige and Skoog medium (16). The seeds were surface sterilized with 70% ethanol for 2 min followed by 0.1% SDS-$HgCl_2$ for 8 min and rinsed using sterile distilled water with six changes. The surface sterilized seeds were then suspended in 0.1% agarose solution and spread evenly on Petri plate containing solid MS medium. To synchronise germination, the plates were placed at 4°C for 48 h and then incubated in the culture room at 22-24°C with photoperiod of 16 h light (150 µmol m^{-2} s^{-1}) and 8 h darkness till the seedlings established themselves. About 4-5 seedlings were transferred by 14 days of plating (4-6 leaf stage) to pots of 6 cm diameter filled with autoclaved Soilrite (Keltech Energies Ltd, Bangalore) and sub-irrigated with Hoagland's solution.

Plant transformation and raising of subsequent generations — The floral dip method for *Agrobacterium*-mediated transformation of *Arabidopsis* was followed (17). The mature seeds harvested were dried, surface sterilized and plated on MS medium containing 250 mg l^{-1} cefotaxime and 50 mg l^{-1} kanamycin. The kanamycin-resistant seedlings (T_1) were transferred to Soilrite. The T_1 lines or primary transformants were grown separately to collect

the T_2 seeds from each line. The T_2 seeds were plated on selection medium with 250 mg l^{-1} cefotaxime and 50 mg l^{-1} kanamycin, and the histochemical GUS assay was carried out at different stages of development of the healthy T_2 plants.

Histochemical GUS assay — The expression pattern of GUS was determined by histochemical GUS assay as described (18). The plants at different stages of growth and various tissues were fully submerged in the GUS assay buffer taken in glass vials and subjected to vacuum infiltration for 5 min. The vials were closed and incubated overnight at 37°C. The GUS assay buffer was carefully decanted and chlorophyll and other pigments were removed by successive addition of 70% ethanol. The expression of GUS was visualized under a Nikon HFX II light stereomicroscope fitted with a fibre optic light source.

RNA isolation and RT-PCR — Total RNA was extracted from petioles, upper-half of leaf lamina and whole leaves of 3-week-old *Arabidopsis* plants using TRIZOL® reagent (Invitrogen). The intactness of isolated RNA was analysed in 1% agarose gel and the quantity of RNA was estimated by using NanoDrop® Spectrophotometer ND1000 as per the specifications and the concentration of RNA obtained in ng μl^{-1}. The RNA samples were treated with DNase I (MBI Fermentas) prior to the RT-reaction to remove the possible contaminating DNA. The RT reactions were performed using One-step RT-PCR Kit (QIAGEN) (13). In order to unambiguously detect the presence of contaminating genomic DNA, primers leading to the specific amplification of genomic DNA were designed flanking the introns of 287 and 107 bp present in the *SAHH1 (HOG1)* and *ACTIN8* genes, respectively. Primers 57RT-1 and 57RT-2 were used to amplify 650 bp *SAHH1* transcripts, while primers ACT8F0 and ACT8R0 were used to amplify the 292 bp *ACTIN8* transcripts (positive control).

In silico analysis of promoter sequences — The nucleotide sequences of intergenic or promoter region and other gene sequences were downloaded from TAIR (The *Arabidopsis* Information Resource) website available at *http://www.arabidopsis.org. In silico* analysis of the 3 Kb *cis*-regulatory region upstream of *SAHH1* of *Arabidopsis,* was performed using PLACE (19). The GeneACT software (20) available online at *http://promoter.colorado.edu/geneact* was used for comparative analysis of the *cis*-regulatory motifs of *SAHH* in rat, mouse and human genomes. Sequence alignment and comparative analysis of *cis*-regulatory regions of *SAHH1* and *SAHH2* were carried out using DiAlignTF of *Genomatix (http://www.genomatix.de/online_help/help_dialign/dialign_TF.html).*

Results and Discussion

The promoter fragments were cloned upstream to promoter-less *GUS* in a binary transformation vector and mobilized into *Agrobacterium* strain GV3101. The expression pattern of GUS was studied in the T_2 lines of transgenic *Arabidopsis* plants.

GUS expression driven by 1.5 Kb fragment (B1.5) upstream to SAHH1 — The 543 bp (B0.5) and 1181 bp (B1.0) fragments upstream to ATG of *SAHH1* gene were capable of driving constitutive GUS expression in the T_2 lines of transgenic *Arabidopsis* (18). The 1 Kb and 1.5 Kb fragments included the 58 bp fragment otherwise deleted in the mutant due to T-DNA insertion. The 1585 bp fragment (B1.5) consisted of 1Kb fragment upstream to ATG of *SAHH1* and the 452 bp (Fig. 1A). In the seven T_2 lines generated for B1.5 construct, GUS expression was very intense in the roots and root tips and expressed in almost all the aerial plant parts. The expression was constitutive and intense in young seedlings as shown in Fig. 1B. The expression pattern of GUS in transgenics carrying B0.5 and B1.0 constructs was also constitutive. Addition of the 452 bp fragment did not alter GUS expression pattern indicating the absence of the *cis*-regulatory motifs in the 452 bp fragment that may otherwise alter or reduce GUS expression. Also, the deleted 58 bp did not contain any negative *cis*-regulatory element for suppressing root-specific expression since the expression in roots was not altered or reduced by including the 58 bp.

cis-regulatory region of SAHH1 — To understand the nature of *cis*-regulatory elements and to delineate the *cis*-regulatory region of the *SAHH1* gene, a comparative histochemical analysis of GUS expression pattern was carried out in B0.5, B1.0 and B1.5 plants. As indicated in Fig. 1B, the GUS expression appeared to progressively increase as the fragment length upstream to *SAHH1* gene increases. *In silico* analysis of the upstream region of *SAHH1* has shown the presence of root specific-motif *as1* at -34 and TATA box at -28 from the annotated TSS of *SAHH1* gene. Expression of GUS was intense in the roots of T_2 progeny of transgenic lines with B0.5 and B1.0. This clearly showed that 543 bp fragment upstream to the *SAHH1* gene had a root-specific motif. A closer examination of the spatial

expression pattern of the reporter gene revealed that the expression of GUS was weak or absent in the petioles of 22-day-old B0.5 plants, while intense expression of GUS was noticed in the petioles of B1.0 plants of the same age (Fig. 1C). Also in some leaves of B0.5, GUS expression was intense in the upper half of leaf lamina. RT-PCR analysis of *SAHH1* gene expression carried out using leaf petioles, upper half of the leaf lamina and whole leaf of 22-day-old wild type *Arabidopsis* plants revealed that *SAHH1* gene was expressed in all these analysed tissues of wild type plants (Fig. 1D). As shown in Fig. 1D, the *SAHH1* gene was expressed in the petioles of the wild type plants (lane 1) while the promoter fragment of 543 bp upstream to *SAHH1* gene did not drive the GUS expression in this region. The results suggested that the *cis*-regulatory sequence of *SAHH1* gene might be upstream to or beyond the 543 bp region.

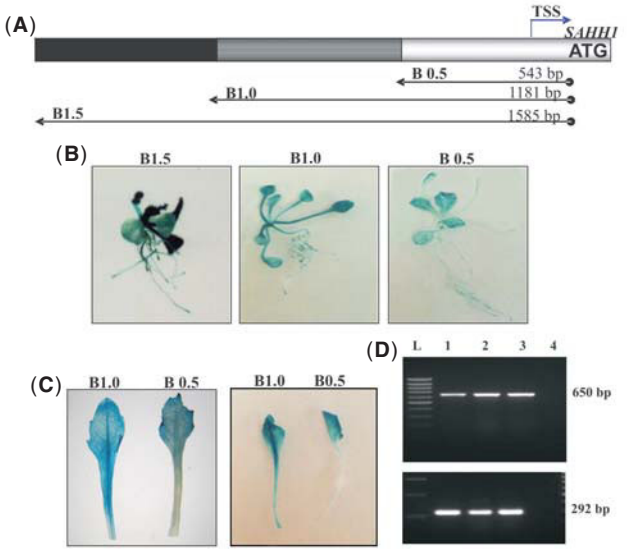

Fig. 1. Cloning of B1.5 (1.5 Kb upstream of *SAHH1* gene) and comparison of GUS expression in the transgenic *Arabidopsis* lines carrying B0.5 and B1.0 fused to promoter-less *GUS*. (**A**) Schematic representation of the fragments B0.5, B1.0 and B1.5 upstream to *SAHH1* gene. The TSS and ATG of *SAHH1* gene are marked. The size of the fragment amplified is given in bp, (**B**) Constitutive GUS expression in 20-day-old plants of B0.5, B1.0 and B1.5 constructs, respectively. Notice the progressively increasing intensity of GUS in transgenics with promoter fragments B0.5, B1.0 and B1.5, (**C**) GUS expression in leaves of 20-day-old B1.0 and B0.5 transgenics. Note the absence of expression in the petioles of B0.5 leaves, and (**D**) RT-PCR analysis of *SAHH1* transcripts in petioles of 22-day-old wild type *Arabidopsis* plants. Lanes 1, 2 and 3 are RT using RNA from petioles, upper half of leaf lamina and whole leaves of 22-day-old wild type *Arabidopsis* plants, respectively. Lane 4 is negative water control and L is 100 bp ladder. Upper panel is RT-PCR of *SAHH1* transcript and lower panel represents RT-PCR of *Actin8* transcripts using same quantity of RNA from the respective tissues.

GUS expression driven by 376 bp promoter fragment of SAHH1 — The minimum length of the promoter fragment upstream to *SAHH1* showing expression in roots was 543 bp (B0.5). In a previous study, the experimental TSS was found to be a C nucleotide located 61 bp upstream to the ATG of *SAHH1* and 8 bp interior to the annotated TSS (14). To determine the nature of expression of the 543 bp promoter fragment, the particular promoter was deleted from both ends (98 bp interior from the 5′ end and 69 bp interior from the 3′ end of the 543 bp) and a 376 bp fragment (SAHD1) upstream to the experimental TSS of *SAHH1* was cloned into pBI101. Schematic representation of these cloned promoter fragments of *SAHH1* is given in Fig. 2A. In all the 3 T_2 lines analysed, a similar expression pattern of GUS was observed in the inflorescence and pods. GUS expression was variable in the leaves and absent from the roots in all the three T_2 lines (data not shown). In 55-day-

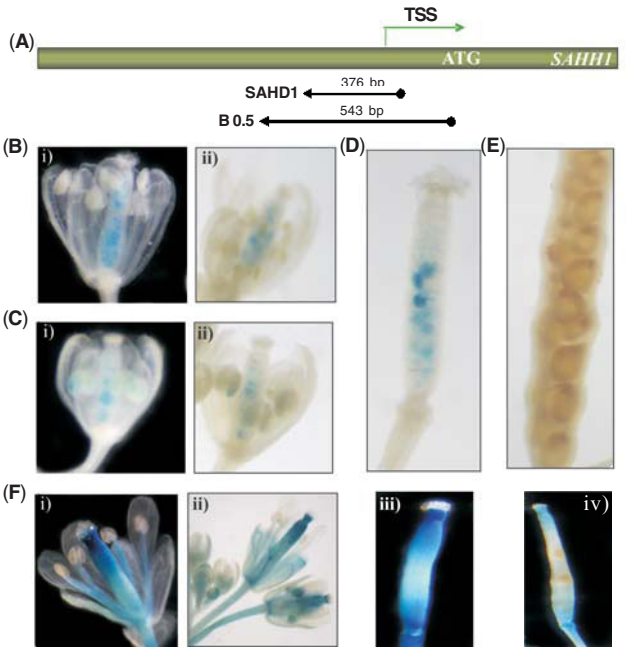

Fig. 2. GUS expression in 7-week-old plants of T2 lines of SAHD1 (376 bp from experimental TSS of *SAHH1* gene). (A) Schematic representation of the fragments cloned upstream to ATG and TSS of *SAHH1* gene, (**B**) i) and ii) GUS expression in young developing ovules in T_2 progeny of transgenic lines of SAHD1. Notice the absence of GUS expression in developing anthers of the same flower, (**C**) i) and ii) GUS expression in developing anthers and developing ovules of the same flower, (**D**) GUS expression in developing ovules of immature silique, (**E**) GUS expression is absent from ovules/seeds of mature silique, and (**F**) GUS expression in i) and ii) flowers, iii) immature silique and iv) mature siliques of T2 progeny of transgenic lines with B0.5 construct fused to promoter-less GUS vector. The GUS expression is nearly constitutive. However, no expression could be detected in the developing ovules.

old plants, GUS expression was present in the developing ovules of flowers and in seeds of developing pods (Fig. 2B, C, D). In certain flower clusters, in addition to GUS expression in developing ovules, faint expression was noticed in the anthers of unopened floral buds and developing flowers, of the same flower cluster or even the same flower (Fig. 2C). GUS expression was prominent in developing seeds of immature siliques (Fig. 2D), while the seeds of mature siliques did not show any expression (Fig. 2E). The 543 bp fragment on the other hand drives nearly constitutive GUS expression in the flower clusters, immature and mature siliques but is absent in developing ovules (Fig. 2F). The 5′ end of the cDNA clone isolated from maize endosperm had similarity to SAHH (21). Also, the *SAHH1* gene has been repoted to be involved in methylation cycles in endosperm development of barley seeds and the activity was found to increase during grain development (10). The locus *SAHH1/HOG1* is also known as EMBRYODEFECTIVE1395 and the homozygous mutation is embryonic lethal while the heterozygotes for *SAHH1* mutation show embryo abortive phenotype (11) suggesting that *SAHH1* is involved in embryonic development. Recently, higher levels of *SAHH* transcripts were found in the developing ovules and microspores of *Arabidopsis*, and SAHH abundance was found to be distinctly higher in seed coats of immature seeds, while adenosine kinase, an associated enzyme in methylation cycle, was hardly detectable. The lack of coordinated expression of the two enzymes involved in maintenance and recycling of SAM-dependent methylation of *Arabidopsis* suggested a non-methyl-related role of *SAHH1* in the seeds (22).

In silico analysis of SAHH1 promoter for seed-specific motifs — *In silico* analysis of the 3 Kb *cis*-regulatory region upstream of *SAHH1* for seed-specific motifs using PLACE revealed the presence of three SEF3 motifs in the 3 Kb region, one of them being located at 165 bp from TSS and which falls within the 376 bp SAHD1 promoter (Fig. 3). Also, three SEF4 motifs were identified but, none in the first 376 bp region. SEF (Soybean Embryo Factor) motifs present in soybean β-conglycinin gene are involved in seed-specific expression. SEF3 binds to a region composed of two elements located at **-**183 to -169 bp and -153 to -134 bp relative to the start of transcription of β-*conglycinin* and the activity increases in embryos during β-conglycinin synthesis and decreases in maturing seeds (23). The putative motifs responsible for seed-specific expression identified in the 3Kb *cis*-regulatory region of *SAHH1* along with the sequences and positions of motifs from the annotated TSS is given in Table 2.

The most dramatic morphological and biochemical events of seed development, such as rapid expansion of the cotyledons, synthesis and accumulation of proteins, lipids, and, in some cases carbohydrates occur during maturation. Maturation also involves significant physiological adaptations that occur within the seed to ensure embryo dormancy, including accumulation of ABA during late embryogenesis to prevent precocious germination prior to desiccation (24). Analysis of the expression pattern of GUS driven by the 376 bp upstream to TSS of *SAHH1* clearly showed GUS expression in developing ovules and anthers but not in mature seeds/anthers. *In silico* analysis of the promoter region of *SAHH1* also has shown the presence of ABRE (ABA Responsive Elements), two of them at -197 and -199 bp from TSS. GUS expression was seen through out the developing ovule and not localized to embryo which points to the involvement of factors other than SEF3 such as Skn-1 motif (GTCAT) located at -110, -320 and -142 bp relative to TSS of *SAHH1* and identified by PlantCARE. All these motifs are located within the 376 bp SAHD1 promoter fragment. The Skn-1 motif is involved in endosperm-specific expression in rice (25). Cereal endosperm is persistent while cellular endosperm of *Arabidopsis* is depleted gradually as the embryo grows and cotyledons take over the support function for the germinating embryo. In mature seeds not yet released from the silique, a massive embryo fills the ovule. This also corroborates with the presence of endosperm-specific motifs because the GUS expression is not detected in mature seeds. The absence of ovule-specific expression in B0.5 promoter fragment indicated the presence of the suppressor/negative regulatory element in the 98 bp fragment at the 5′ end of B0.5 (543 bp) other wise deleted from SAHD1 of negative regulatory effect of elements in the 69 bp 5′ UTR of *SAHH1*. Also, GUS expression in developing ovules as well as anthers/microspores, as observed in this experiment, supplements the observations of earlier report (22) on the expression of *SAHH1* indicating the relatively higher activity of *SAHH1* in ovules and anthers. The *cis*-regulatory motif of *SAHH1* responsible for GUS expression in developing ovule/endosperm as well as anthers/microspores were also detected in the 376 bp fragment upstream to the TSS of *SAHH1* gene.

In silico comparative analysis of SAHH promoter region — *In silico* comparative analysis of the 3 Kb region upstream of *SAHH1* gene in various organisms, such as rat, mouse and human was carried out using GeneACT software. A

Table 2. Putative *cis*-regulatory motifs associated with the seed-specific expression of *SAHH* gene showing the position relative to the TSS in bp

Motif	Organism	Consensus sequence	Position upstream to TSS
SEF3	*Arabidopsis*	AACCCA	-160 to -165,-956 to-961, -2712 to-2717
SEF4			
1	*Arabidopsis*	RTTTTTR	-966 to -972,-2324 to -2330, -2734 to -2740
2	Rat	ATTTTTG	-195 to -201, -202 to -208
3	Mouse	GTTTTTG	-2246 to -2252
4	Human	ATTTTTA	-524 to -530, -1689 to -1695, -1836 to -1840, -2021 to -2027, -2602 to -2608, -2610 to -2616
ABRELATERD1	*Arabidopsis*	ACGTG	-193 to -197
ABRERATCAL	*Arabidopsis*	MACGYGB	-194 to -198, -1083 to -1087

Note: R is A/G, M is A/C and Y is C/T

comparative analysis of the *cis*-regulatory region revealed that the soybean embryo factor4 (SEF4) for seed-specific expression was conserved in orthologous systems (Fig. 3). The functional conservation of *cis*-regulatory elements in organisms which are distantly related or which share the least common ancestry (greater evolutionary divergence) would reveal the modules which are functionally conserved along the process of evolution and thus preserved in the genome for carrying out certain essential and critical functions. Regulatory elements in untranslated regions might be detected by sequence comparison or "phylogenetic footprinting" (26). Earlier, a functional dissection of the *apetala3* promoter region in *A. thaliana,* was carried out along with the identification of conserved regions of promoter sequence between *A. thaliana* and *B. oleracea* (27). Analysis of inter-specific patterns of sequence conservation in promoter regions of two genes, *chs* (chalcone synthase) and *apetala3,* showed

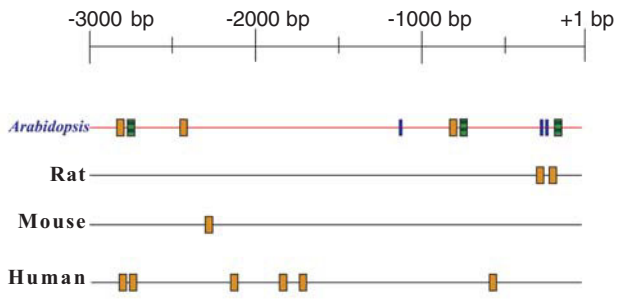

Fig. 3. Comparative analysis of 3 Kb *cis*-regulatory region upstream of TSS of *SAHH* gene promoter in orthologous systems showing conserved motifs associated with seed-specific expression. The SEF4 motif conserved across orthologous systems is shown as shaded box, the SEF3 motif in cross-hatched box and ABRE (ABA Responsive Elements) in vertical lines.

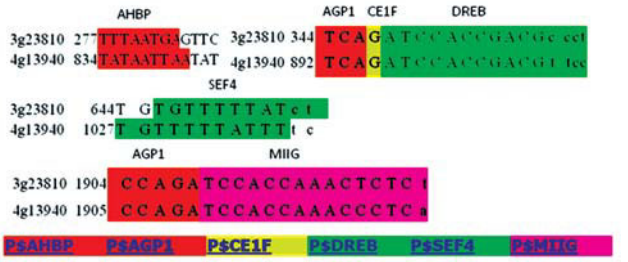

Fig. 4. Comparative analysis of *cis*-regulatory regions (2Kb upstream of annotated TSS) of *SAHH2* and *SAHH1* genes in *Arabidopsis* using DiAlignTF showing SEF4 motif. 3g23810 represents *cis*-regulatory sequence of *SAHH2* gene and 4g13940 is that of *SAHH1* gene. The numbers indicate the position of the base shown in bp from the 5′ end of the 2 Kb sequence. The transcription factor binding sites are shown in boxes with the names above the boxes as AHBP, AGP1, CE1F, DREB, SEF4 and MIIG.

that known functionally important regulatory elements were conserved among crucifer relatives of *Arabidopsis* (28).

A comparative analysis of the upstream 2 Kb *cis*-regulatory region between *SAHH1* and the paralogue, *SAHH2* showed the presence of several putative motifs which were related to the function of these genes and threw light on the possible functional conservation. The six conserved motifs are highlighted in Fig. 4. AHBP is binding site for *Arabidopsis* Homeobox Protein, AGP1 for binding of AG motif Binding Protein1 which is a plant GATA-type zinc finger protein binding to AG motif (AGATCCAA).The other important motifs identified in the region were CE1F (Coupling Element1 Factor), SBox (Sugar and ABA-responsive elements), DREB (Dehydration Responsive Element Binding), SEF4 (Soybean Embryo Factor4 binding) and MIIG (MYB IIG binding). *SAHH1* is dehydration-responsive gene (29,30) and the conservation of DREB in both paralogues suggested that both might be responding

to salinity or dehydration stress. The SEF4 motif found in *SAHH1* was also conserved in the paralogue *SAHH2* suggesting the importance of both the genes in seed development. *Arabidopsis* is a cryptic tetraploid and the genome is a composite of the two ancestral genomes (31). The functional divergence between duplicates is required for their long-term retention in a genome. One gene may maintain the ancestor's function and the other may freely mutate and evolve. The fate of duplicates may be neo-functionalisation (gain of function) and pseudogenisation or non-functionalisation or simply loss of function as proposed (32). In lucerne, one of the two SAHH enzymes or isozymes was shown to express at a low level in association with the house-keeping activities, while another was found to be expressed in vascular tissues (7). Similarly, in barley, one of the SAHH enzymes, HvAHH1, was expressed nearly in all tissues and had house-keeping function while the other (HvAHH2) exhibited a preference of expression for seeds (10).The *SAHH2* of *Arabidopsis* must have lost the property of supporting life at all stages after polyploidisation (11). The *SAHH1* and *SAHH2* genes might be sharing certain functions, one of them being the involvement of both genes in seed development, although the mutation in *SAHH2* is not reported to affect the female fertility as that of *SAHH1* (11).

Thus, the conservation of ovule/seed-specific motifs of *SAHH1* in the orthologous systems as well as paralogue (*SAHH2*) described in the present investigation suggests the significance of the enzyme, S-adenosyl homocysteine hydrolase in the embryo/seed development. In addition, this study has led to the identification of a seed-specific motif within the 376 bp upstream to the experimental TSS of *SAHH1.*

Acknowledgements

We thank NATP (CGP-II/253), ICAR for the financial assistance. TPS and CS acknowledge the Fellowships by CSIR. The assistance provided by the staff of the National Phytotron Facility, IARI, New Delhi for growing *Arabidopsis* is gratefully acknowledged.

Received 31 March, 2008; accepted 23 August, 2008.

Online published 3 October, 2008.

References

1 **Stepkowski T, Brzeziński K, Legocki AB, Jaskólski M & Béna G,** *Mol Phylogenet Evol,* **34** (2005) 15.

2 **Mull L, Ebbs ML & Bender JA,** *Genetics*, **174** (2006) 1161.

3 **Kawalleck P, Plesch G, Hahlbrock K, & Somssich IE,** *Proc Natl Acad Sci,* USA, **89** (1992) 4713.

4 **Tanaka H, Masuta C, Kataoka J, Kuwata S, Koiwai A & Noma M,** *Plant Sci*, **113** (1996) 167.

5 **Weretilnyk EA, Alexander KJ, Drebenstedt M, Snider JD, Summers PS & Moffatt BA**, *Plant Physiol,* **125** (2001) 856.

6 **Zheng J, Zhao J, Tao Y & Fu J,** *Plant Mol Biol*, **55** (2004) 807.

7 **Abrahams S, Hayes CM & Watson JM,** *Plant Mol Biol,* **27** (1995) 513.

8 **Medina-Suarez R, Manning K, Fletcher J, Aked J, Bird CR & Seymour GB,** *Plant Physiol,* **115** (1997) 453.

9 **Hadfield KA, Dang T, Guis M, Pech J-C, Bouzayen M & Bennett AB,** *Plant Physiol*, **122** (2000) 977.

10 **Radchuk VV, Sreenivasulu N, Radchuk RI, Wobus U & Weschke W,** *Plant Mol Biol,* **59** (2005) 289.

11 **Rocha PSCF, Sheikh M, Melchiorre R, Fagard M, Bouter S, Loach R, Moffatt B, Wagner C, Vaucheret H & Furner I,** *Plant Cell*, **17** (2005) 404.

12 **Li C-H, Jiang S-M, Wang L-J & Chen X-Y,** TAIR accession Publication:501721780, 2007.

13 **Sivanandan C***, Ph.D. thesis*, Indian Agricultural Research Institute, (2004).

14 **Sivanandan C, Sujatha TP, Anand Mohan Prasad, Resminath R, Thakare, DR, Bhat SR & Srinivasan,** *Biochim Biophys Acta,* **1731** (2005) 202.

15 **Sambrook J & Russell DW,** *Molecular cloning: A laboratory manual* (3-Volume Set), Cold Spring Harbor Press, USA, (2001).

16 **Murashige T & Skoog F,** *Physiol Plant*, **15** (1962) 473.

17 **Clough S & Bent A,** *Plant J,* **16** (1998) 735.

18 **Jefferson RA, Kavanagh TA & Bevan MW,** *EMBO J, 6* (1987) *3901.*

19 **Higo K, Ugawa Y, Iwamoto M & Korenaga T,** *Nucleic Acids Res,* **27** (1999) 297.

20 **Cheung TH, Kwan YL, Hamady M & Liu X,** *Genome Biol,* **7** (2006) R97.

21 **Shen B. Carneiro N, Torres-Jerez I, Stevenson R, McCreery T, Helentjaris T, Baysdorfer C, Almira E, Ferl R, Habben J & Larkins B,** *Plant Mol Biol,* **26** (1994) 1085.

22 **Pereira LAR, Todorova M, Cai X,. Makaroff CA, Emery RJN & Moffatt BA,** ***J Exp Bot,*** **58** (2007) 1083.

23 **Allen RD, Bernier F, Lessard, PA & Beachy RN,** *Plant Cell,* **1** (1989) 623.

24 **Thomas TL,** *Plant Cell*, **5** (1993) 1401.

25 **Washida H, Wu CY, Suzuki A, Yamonouchi U, Akihama T, Harada K & Takaiwa F,** *Plant Mol Biol,* **40** (1999) 1.

26 **Wasserman WW, Palumbo M, Thompson W, Fickett JW & Lawrence CE,** *Nat Genet,* **26** (2000) 225.

27 **Hill TA, Day CD, Zondlo SC, Thackeray AG, Irish VF,** *Development*, **125** (1998) 1711.

28 **Koch MA, Weisshaar B, Kroymann J, Haubold B & Mitchell-Olds T,** *Mol Biol Evol*, **18** (2001) 1882.

29 **Weretilnyk EA, Alexander KJ, Drebenstedt M, Snider JD, Summers PS & Moffatt BA,** *Plant Physiol,* **125** (2001) 856.

30 **Zheng J, Zhao J, Tao Y & Fu J,** *Plant Mol Biol*, **55** (2004) 807.

31 ***Arabidopsis* Genome Initiative,** *Nature,* **408** (2000) 796.

32 **Ohno S,** *Evolution by gene duplication,* Springer-Verlag, New York, (1970).