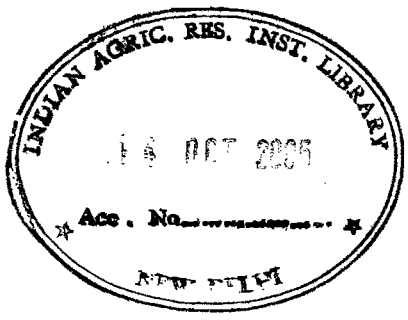
TAGGING, ISOLATION AND CHARACTERIZATION OF A ROOT-SPECIFIC PROMOTER FROM *ARABIDOPSIS*

C. SIVANANDAN



NATIONAL RESEARCH CENTER ON PLANT BIOTECHNOLOGY
INDIAN AGRICULTURAL RESEARCH INSTITUTE
NEW DELHI- 110012
2004

TAGGING, ISOLATION AND CHARACTERIZATION OF A ROOT-SPECIFIC PROMOTER FROM *ARABIDOPSIS*

By

C. SIVANANDAN

A thesis submitted to the Post Graduate School,
Indian Agricultural Research Institute, New Delhi
in partial fulfillment of requirements for the award of degree of

DOCTOR OF PHILOSOPHY

IN

MOLECULAR BIOLOGY AND BIOTECHNOLOGY

Approved by

Chairman

Srinivasan

Co-Chairman

K.R. Koundal

Members

I.M. Santha

K.V. Prabhu

Dr. Srinivasan
Principal Scientist

NRCPB, IARI
New Delhi
110 012

CERTIFICATE

This is to certify that the thesis entitled "**Tagging, Isolation and Characterization of a Root-Specific Promoter from *Arabidopsis***", submitted to the Faculty of Post Graduate School, Indian Agricultural Research Institute, New Delhi, in partial fulfillment of the requirements for the award of Doctor of Philosophy in Molecular Biology and Biotechnology by C. Sivanandan embodies the results of bonafide research work carried out by him under my supervision. He has submitted no part of the thesis for any other degree or diploma. He has also duly acknowledged all the assistance received by him during the course of the work.

Srinivasan

Place: New Delhi
Date: 22nd June 2004

(**Srinivasan**)
Chairman
Advisory Committee

ACKNOWLEDGEMENTS

I am highly indebted to Dr. Srinivasan, Principal Scientist, NRCPB, IARI my chairman for the ample freedom that he has always given me. His critical judgments, timely and valuable advice and his sense of humor have always inspired me.

I also acknowledge the constant encouragement provided by Prof. K.R. Koundal, Project Director, NRCPB, IARI, who also happened to be my co-chairman.

I would like to extend my thanks to Dr. I.M. Santha, member, advisory committee, for her encouragement and constructive criticism. I would never forget the help rendered by Dr. K.V. Prabhu, member advisory committee for providing us with uninterrupted access to the National Phytotron Facility and for thoroughly reviewing my manuscript. I enjoyed the several discussions that I had with Dr. S.R. Bhat, Co-PI in the NATP project in our laboratory, I reserve a very big THANKS for him. I salute Dr. S.K. Raina, Professor NRCPB for his energetic smile constantly egging me on.

I acknowledge the help rendered to us by Dr. J.P. Khurana, Delhi University South campus by gifting us Arabidopsis seeds.

I also extend my sincere thanks to Dr. P.K. Jain and Mr. R.S. Jaat, scientists in our laboratory for their guidance and companionship. I consider myself lucky to have got a wonderful set of lab mates in Resminath, Anand Mohan and Dhiraj. I still cherish some of the memorable moments I had with them. I also gratefully acknowledge all my friends, Gundu, Gow, Deepa, Abhijit, Deepti, Ravi, Sudhir, M. Anil Mehta (mon professeur de Francaise), Sekhar, Selvi, Praks, Kums, Ajai ,Babu, P.S, Deepak, Elayaraja, my former batchmates of Coimbatore and numerous other people especially Arul and Raghuraman who strived to keep me in good spirits.

I also thank Vedji for all his help in the lab and Mr. A.K. Jain for his help in completing the academic formalities in time.

I can't, but help mention the emotional support and encouragement given to me by my amma, appa, all my annas and my beloved wife, Anju who never wanted her name to appear on this page. I owe them more then anything else in the world. I specially thank Chittassi, Sasiappan and Appu for their companionship.

I acknowledge the financial assistance provided by CSIR-UGC in the form of JRF/SRF grant during the course of this work. I also immensely benefited from the funds granted to the laboratory vide the NATP project (26-03).

(C. SIVANANDAN)

CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Identifying and understanding DNA sequences involved in various developmental processes of plants not only serve academic interest but also is critical for devising intelligent strategies for genetic engineering. With the increasing application of transgenics in crop improvement, the importance of a variety of regulatory elements that determine temporal and spatial expression pattern of the associated gene is self-evident. Future strategies to improve crop plants will significantly benefit from the isolation and characterization of novel genes and regulatory sequences.

A variety of approaches are used for cloning and functional identification of a gene. Amongst them, an important approach is T-DNA insertional mutagenesis. Gene trap vectors based on T-DNA, have been effectively used to identify promoters and genes with interesting expression patterns (Wei *et al.*, 1997; De Greve *et al.*, 2001 and Mollier *et al.*, 2000). The first step in this process is to generate a mutagenized population carrying random insertions, using gene/promoter traps. Gene tagging and promoter trapping strategy relies on the transfer and random integration of a promoter-less reporter gene (usually *gus*) into the plant genome. The expression of the reporter gene, which could be activated by adjacent regulatory elements, is then dependent on the location of the promoter-less reporter gene in the plant genome.

A clear advantage of these gene-trapping systems is that the expression pattern of the tagged gene can be studied in detail by analyzing GUS staining pattern during the development of the plant under a variety of

environmental conditions. Such a detailed knowledge of the expression pattern can be very helpful in locating and isolating specific regulatory elements.

Currently, the most ideal system for plant molecular genetic studies is *Arabidopsis* because of its short life span, high fecundity, a small genome size and less repetitive sequences. An easy and efficient *in planta* transformation makes it possible to generate large populations of T-DNA insertion lines without recourse to tissue culture thereby avoiding the tissue culture related problems of somaclonal variation. In addition, the availability of the complete genome sequence of *Arabidopsis* makes the job of identifying and characterizing the functional aspects of genes and their regulatory sequences much easier. There has been a remarkable conservation of elements controlling expression specificity between species. Thus, promoter elements directing tissue specific expression, identified in *Arabidopsis* could be employed in genetic engineering of a variety of crop plants.

Roots are extremely important to the plant, as they are the sites of nutrient absorption and interaction with various microorganisms in the soil. A gene expression map of different genes expressed in roots during various developmental stages has been constructed (Birnboim *et al.*, 2003). An *Arabidopsis* MADS box gene that induces proliferation of lateral roots in response to low levels of nitrate has been cloned (Zhang and Forde, 1998). Attempts are being made to introduce two root-specific genes, MtPT1 and MtPT2, up regulated in response to phosphate starvation into *Medicago truncatula* (Xiao *et al.*, 2003). A promoter element

directing strong root specific expression could be put to a number of biotechnological applications such as enhanced expression of pathogen resistance gene (Jach *et al.*, 1995). A very important role for root specific promoters could be in phytoremediation, i.e. the cleansing up soil of heavy metals and other pollutants using plants (Gleba *et al.*, 1999) and molecular farming (Drake *et al.*, 2003). Thus in the present study, an attempt has been made to tag, isolate and characterize a root-specific promoter from *Arabidopsis* by promoter trapping with the following objectives:

1. Construction of a promoter trap vector with a promoter-less *gus* gene.

2. Generation of a population of T-DNA mutagenized *Arabidopsis* plants, using the promoter trap vector.

3. Identification of mutants showing root specific expression of GUS.

4. Isolation of the flanking sequences and characterizing the upstream regulatory sequences from those mutants.

2.1 EUKARYOTIC PROMOTERS

The promoter is the key *cis*- acting regulatory region generally found at the 5' end upstream of a gene controlling its transcription. In the eukaryotes they conform to a basic structural plan (Fig. 1).

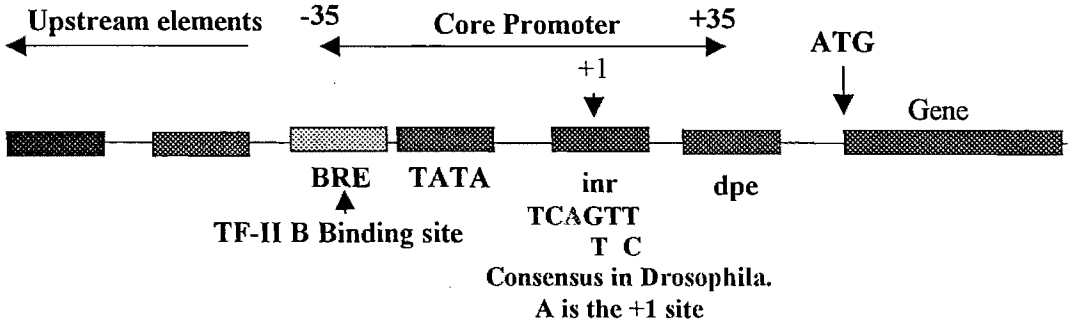


Fig 1: Structure of a typical eukaryotic promoter

Though the arrangements of different sequence motifs vary, the core promoter sequence (located between 35 nt upstream of transcription start site and 35 nt downstream of transcription start site) more or less remains the same. The most common core promoter motif to be identified is the TATA box (Goldberg, 1979 and Breathnach and Chambon, 1981) which is mostly located 25-30 bases upstream of the transcription start site. The consensus sequence of TATA box is TATAAA, however a wide range of sequences can perform the same function (Singer *et al.*, 1990). The next important motif, the initiator element (inr) that encompasses the transcription initiation site has been identified in several eukaryotes. Another important motif is the downstream promoter element (dpe) located 28 to 32 bases downstream of the initiation site. These sites are critical

because several transcription factors, DNA binding proteins, interact with these and help in initiating transcription. These core promoters occur in virtually all genes. Due to this fact they were never thought to contribute to regulation of transcription. However the picture has now changed as more and more promoters are being isolated. The extensive similarity in different motifs has vanished and given rise to a more complicated picture. There are several promoters, which do not have even TATA boxes. It was estimated that approximately 43% of 205 core promoters in *Drosophila* contain a TATA box (Kutach and Kadonaga, 2000). In humans, it was found that about 32% of 1031 potential promoter regions contain a putative TATA box motif (Suzuki *et al.*, 2001). In plants, the tobacco *psa*DB gene is the first example of a TATA-/Inr+ promoter (Nakamura *et al.*, 2002). Of 232 plant promoter sequences analyzed using a computer algorithm designed for identifying *cis*-acting regulatory elements (Tsunoda and Takagi, 1999), 41 were found to be without a TATA box motif and a majority of these genes were involved in photosynthesis. The presence of two TATA boxes synergistically regulating transcription of tubulin genes in soybean (Doyle and Han, 2001) and three TATA boxes responsible for the efficient initiation of β-phaseolin gene in *Phaseolus vulgaris* (Grace *et al.*, 2004) underline the importance, complexity and variation of core promoter in the organization of transcriptional regulation. Even transcription, in *Drosophila* more commonly initiates at a cluster of multiple sites in the vicinity of the inr (and not necessarily at the same position vis a vis an inr consensus sequence as depicted in fig. 1). Thus, although enhancer elements that are located several bases upstream of the promoter itself are key to

transcriptional gene regulation, the role of core promoter is critical. The organization and regulation of promoter elements in eukaryotes has been comprehensively reviewed (Butler and Kadonaga, 2002 and Smale, 2001). Eukaryotic cells carry information for a very large number of complex processes involved in maintaining life. In the genomic era, it is being realized that the gene content of several organisms is much less than previously thought. In the draft human genome sequence, compared to the previously thought 1,00,000 genes, only about 30000 genes were counted. Even in *Arabidopsis*, the first plant genome to be sequenced (*Arabidopsis* Genome Initiative, 2000) the number of protein coding genes is only 26000. The same is true for *Drosophila* (Hild *et al.*, 2003). Answers to questions like, how are these organisms able to sustain such a tremendous number of life processes with a limited number of genes, hold the key to our understanding of genetic regulation.

2.2 CLONING OF PROMOTERS

Timing and levels of gene expression are crucial parameters for proper development of a eukaryote. Gene expression pattern can be regulated at different levels and can involve the specific action of a plethora of *cis* and *trans*-acting factors before and during transcription as well as a number of post-transcriptional modification processes. The regulation of gene expression is manifested in the form of tissue-specificity, developmental-specificity and a wide range of expression levels in response to endogenous and exogenous stimuli. Genetic dissection of the 5' upstream sequences are being carried out to identify DNA sequences involved in controlling gene expression since long (For example Van

Haaren and Houck, 1991; Yamamoto *et al.*, 1995; Sjodahl *et al.*, 1995 and Meister *et al.*, 2004). In tomato, functional maps of regulated promoters are also available, specifically delineating the contribution of individual *cis*-acting elements towards the observed phenotype (Van Haaren and Houck, 1993; Montgomery *et al.*, 1993a and b; Xu *et al.*, 1996; Santino *et al.*, 1997 and Deikman *et al.*, 1998). In order to understand gene regulation, cloning of upstream regulatory regions are imperative. Early methods to isolate promoters revolved around identifying a gene from a genomic DNA library using its mRNA or cDNA clone leading to cloning the corresponding genomic clone and then their regulatory elements. Some of these approaches have led to the isolation of promoter elements regulating cell and tissue specific expression (Koltunow *et al.*, 1990; Albani *et al.*, 1992). In recent years, more powerful techniques like subtractive hybridization, differential screening and differential display etc. have been used to isolate their respective genomic clones (Feng *et al.*, 2004; Wang *et al.*, 2003; Caliskan *et al.*, 2003 and Matsuyama *et al.*, 1999). ESTs from a specific tissue can also be screened to isolate novel genes. The seed company, Pioneer Hi Bred International is using large scale single pass sequencing to screen the maize genome database to identify novel promoter candidates especially those which are specific to seeds and endosperm (http://abstracts.aspb.org/aspb1998/43/0069.shtml)

PCR based strategies have been used to clone upstream regulatory regions. A procedure involving restriction of genomic DNA followed by ligation of an adapter to the 5' end and then performing PCR with a gene specific and an adapter specific primer, adapter mediated PCR (see for

example Ukai *et al.*, 2002) has been followed to clone fruit and seed specific promoters (Anjanasree, 2003 and Anandhan, 2004). Adapter PCR can be used on genomic DNA itself. RT-PCR on isolated RNA followed by 5' RACE (Random Amplification of cDNA ends) leads to the cloning of the complete gene (Brenner *et al.*, 1996 and Elrouby and Bureau, 2000). Once the complete gene is identified and the transcription start site known, it is very easy to clone the promoter of the gene using various PCR based techniques. However the most versatile strategy to isolate tissue-specific promoters is the promoter trapping technique by T-DNA (Covered in detail in section 2.4.3) or transposons in plants. Promoter trapping uses T-DNA of *Agrobacterium* or transposons as random mutagens. In promoter trapping by T-DNA tagging, the binary vector used has a promoter-less reporter gene. Promoter trapping (e.g. Lindsey *et al.*, 1993 and Mollier *et al.*, 1995) coupled with an ideal reporter gene e.g. β glucuronidase (*gus*-A) (Jefferson *et al.*, 1987) is very convenient because one can directly visualize the reporter gene expression in the tissue of interest. The upstream sequence flanking the T-DNA insert can be isolated by inverse PCR (Lindsey *et al.*, 1993), TAIL PCR (Liu *et al.*, 1995) or by plasmid rescue (e.g. Nakazawa *et al.*, 2001) from the mutant plant exhibiting GUS expression in the tissue of interest. T-DNA insertional mutagenesis in *Arabidopsis* is now routine and populations generated by using such vectors find extensive use in functional genomics.

2.3 *ARABIDOPSIS*: A MODEL PLANT

Arabidopsis is an ideal plant for molecular genetic studies. It has a rapid life cycle (about 6 weeks from germination to mature seed), is

amenable to prolific seed production and easy cultivation in restricted space. Further, its small genome (114.5 Mb/125 Mb total) has been completely sequenced (AGI, 2000). Extensive genetic and physical maps of all 5 chromosomes are available, easily accessible on a website dedicated to the resources related to *Arabidopsis* (www.arabidopsis.org). A large number of mutant lines are also available in stock centers (refer http://www.arabidopsis.org/links/atlinks.jsp). A multinational research community of academic, government and industry laboratories performs research related to *Arabidopsis*. Other than *Antirhinum* and *Petunia* it is the major experimental plant for floral developmental genetics in plants (Ferraraio *et al.*, 2003; Saedler *et al.*, 2001; Pelaz *et al.*, 2001 and Honma and Goto, 2001). It is also a model for embryogenesis although it suffers from the disadvantage of having very small seeds. An exhaustive list of different embryo defective mutations has been prepared (Meinke *et al.*, 2003). The availability of complete genome sequence of *Arabidopsis* (Arabidopsis Genome Initiative, 2000) is likely to broaden and accelerate further research in plant sciences. However, DNA and protein sequences alone fail to provide much biologically relevant information. More than 50% of the predicted *Arabidopsis* genes have not been assigned any specific function. Comparison with sequences in the databases is the simplest way to obtain functional information. *In silico* analysis by itself is only indicative and is generally not sufficient to define the function of a gene. Even in those cases where some indications become available from *in silico* analyses, experimental evidences are required to conclusively prove the predictions. One of the most challenging tasks before plant scientists is

assigning ***biological*** functions to a large number of plant genes. Out of the ~26000 genes identified in *Arabidopsis*, the functions of only a few thousand have been defined with great confidence (Bouche and Bouchez, 2001).

2.4 T-DNA INSERTIONAL MUTAGENESIS IN *ARABIDOPSIS*

Isolation and characterization of mutants is a powerful tool to study genetic and developmental processes in plants. Gene tagging is one of the methods to identify genes known only by their phenotype (Kleckner *et al.*, 1975 and Berg and Berg, 1987). This approach is based on the inactivation of a gene via insertion of a known DNA fragment. The integration site is marked and it is possible to isolate adjacent DNA sequences via this molecular marker. The most commonly used tags are transposons and transforming DNA. Endogenous transposons have long been used for gene tagging in *Antirrhinum majus* and *Zea mays* (Martienssen *et al.*, 1989 and Sommer *et al.*, 1990). The idea of using T-DNA of *Agrobacterium tumefaciens* to clone several genes in *Arabidopsis* was initiated years ago (Herman and Marks, 1989; Koncz *et al.*, 1990 and Yanofsky *et al.*, 1990). Efficient transformation methods utilizing *Agrobacterium tumefaciens* are already available. The original root explant method (Valvekans *et al.*, 1988) allowed one to isolate many hundreds and possibly thousands of transformed plants. A laborious tissue culture process was involved in this method. However, several improvements in plant transformation procedure have been reported. The seed transformation protocol (Feldmann and Marks, 1987) excludes tissue culture induced variations. This protocol has been used to develop collections of *Arabidopsis* transformed lines

(Feldmann and Marks 1987; Feldmann, 1991 and Forsthoefel *et al.*, 1992). The populations so developed have been the source of mutants for different groups working on gene isolation by T-DNA tagging since then (e.g., Marks and Feldmann, 1989; Ross *et al.*, 1997; and Feldmann *et al.*, 1997). An improvement to this seed transformation technique was vacuum infiltration where the whole plant was uprooted and infiltrated with a solution of *Agrobacterium* in vacuum (Bechtold *et al.*, 1993). The vacuum infiltration technique was simple and reliable. The elimination of tissue culture and regeneration greatly reduced hands–on time and even non-experts could achieve successful transformation (Bechtold *et al.*, 1993 and Bent and Clough, 1998). However Clough and Bent, (1998) further refined the technique by simply dipping *Arabidopsis* inflorescences in a suspension of *Agrobacterium tumefaciens*, carrying the desired plasmid. In this floral dip method, each transformed seed of a plant would represent an independent transformation event. This procedure has made it possible to obtain thousands of independent transformants from just a few pots of *Arabidopsis* plants.

An advantage of using T-DNA as the insertional mutagen as compared to transposons is that the T-DNA insertions do not transpose subsequent to insertion and are chemically and physically stable through multiple generations. The T-DNA not only disrupts the expression of the gene into which it is inserted, but also acts as a marker for subsequent identification of the mutation. Since *Arabidopsis* introns are small and because there is comparatively less intergenic material, the insertion of a piece of T-DNA of the order of 5 to 25kb in length generally produces a

disruption of gene function. If a large enough population of T-DNA transformed lines is generated, there are reasonably good chances of finding a transgenic plant carrying a T-DNA insert within any gene of interest. Mutations that are homozygous lethal can also be obtained and maintained in the population in the form of heterozygous plants.

Low copy number and random nature of insertions are some of the other advantages of T-DNA induced mutagenesis. After segregation analysis of a large number of transformants, it was concluded that the average number of independent inserts is 1.5 per diploid genome with 57% of the transformants containing a single insert and 25% containing two (Feldman, 1991). The insertions appeared to be random and no obvious hotspots for insertion of T-DNA could be detected (Azpiroz-Leehan and Feldmann, 1997). However more recent research is proving that these observations could be way off the mark as is evident from genome wide analysis of the distribution of integration sites conducted by Alonso and co-workers (Alonso *et al.*, 2003). A T-DNA mutagenized population of *Arabidopsis* revealed the existence of a large integration site bias at both chromosome and gene levels. A study conducted at the Salk institute of biological studies, La Jolla CA, USA has generated over 225000 independent T-DNA insertions representing near saturation of the entire gene space. The precise locations of 88000 insertions were also identified with mutation in ~73% of predicted *Arabidopsis* genes. The number of integrations was reported to decrease dramatically from the gene rich chromosome arms towards the centromeres. At the gene level a significant bias was observed against integration events in exons and introns in favor

of 5'UTR, 3' UTRs and promoters. The German Plant Genomics Research programme GABI-KAT has also made a similar observation (Rosso *et al.*, 2003).

2.4.1 Getting to the T-DNA flanking sequences

A variety of strategies have been devised to generate and isolate mutants in known genes of *Arabidopsis* by T-DNA or transposon insertional mutagenesis. In these methods, large populations of tagged mutants are generated, which can then be screened for insertions in specific genes. Alternatively, the insertion tags can be individually sequenced and compiled in databases, which can be searched for a disruption event in the gene of interest.

A highly efficient procedure for obtaining mutants in genes identified in sequencing programs takes advantage of the availability of large collections of plants mutagenized by an insertion element. This procedure makes use of the specificity and sensitivity of the PCR reaction to screen for insertions within regions of interest in a large population of mutagenized plant lines. The sensitivity of the PCR is so high that it is possible to detect such an event in large pools (upto a few thousand) of mutagenized plants (Bouchez and Hofte, 1998). The pool is repeatedly subdivided until a single plant carrying the desired insertion is identified. The identified mutant plants are then tested for phenotypes that are predicted to result from loss of function of the gene. A reverse genetics based approach for identification of T-DNA insertions within any sequenced gene in *Arabidopsis* using gene specific, non-degenerate primers has been reported earlier (Krysan *et al.*, 1996). The process will become even easier in the future as genomic

sequence of many more insertion sites will become available. Large collections of the T-DNA insertion lines are being provided to the *Arabidopsis* stock centers at Ohio State University (USA) and Nottingham (UK) by the individuals who constructed them (Table 1). More than 175,000 T-DNA insertion lines of various types are already available from the *Arabidopsis* Biological Resource Center (ABRC). For a fee, the Arabidopsis Knockout facility offers services to screen for mutations in a particular gene (http://www.biotech.wise.edu/arabidopsis). However in the year 2003 a new resource for knockout mutants has become available. The Salk institute, through the work of Ecker and coworkers has established a computer database that contains the precise genomic locations of over 50000 T-DNA insertions (http://signal.salk.edu/cgi-bin/tdnaexpress). By performing a simple BLAST search using the gene of interest, it can quickly be determined if the SALK collection contains a mutation in that gene.

Table 1. Publicly available collections of Arabidopsis T-DNA lines*

Collection	Background	Spl. Features of Vector	Size	Pool size
Feldmann	Ws	-	10,500	100/20/10
Jack	Columbia	GUS enhancer trap	11,300	100/10
INRA	Ws	GUS enhancer trap	3900	100/20
Weigel	Columbia	Activation tag	22,600	100/20/10
Alonso/Crosby/ Ecker	Columbia	-	>20,000	100/10
Scheible/ Souresville	Columbia	Activation tag	~ 60,000	100-350
Bressan/Zhu	C_{24}	Activation tag	79,700	100
Sussmann/ Amasino	Ws	-	37,800	225

*Weigel and Glazebrook 2002

Several groups are working to isolate and sequence DNA fragments flanking insertion sites and are depositing sequences in databases. When these databases reach sufficient size, it will be, possible to find an insertion in any gene of interest simply by consulting the databases and ordering seed from the stock centre. Several groups have initiated systematic sequencing of insertion sites in various populations (Parinov *et al.*, 1999, Speulman *et al.*, 1999, Tissier *et al.*, 1999). In the next few years, it can be expected that the combined size of these populations will cover the systematic disruption of all *Arabidopsis* genes, and that the search for knockout mutants will be greatly facilitated by the use of knockout databases.

Surprisingly, although PCR based reverse genetic screens have been available to *Arabidopsis* researchers, for several years, relatively few informative knockouts obtained in *Arabidopsis* have given rise to a visible, directly scorable phenotypes. Several reports have shown that lack of phenotypes is presumably because of partial and complete functional redundancy besides the ability of higher plants to adapt their physiology to various stresses and constraints without undergoing morphological changes, and by our inability to detect slight physiological alterations and/or weak reductions in fitness (Bouche and Bouchez, 2001).

In order to identify conditional phenotypes, the mutants have to be tested under a wide range of environmental conditions (Meissner *et al.*, 1999). This proposition makes sense, as it is likely that plants, as sessile organisms, have evolved many adaptive traits that allow them to

cope with profound changes in their biotic and abiotic environment. Many loss-of-function mutants are therefore expected to be conditional and revealed only by a specific combination of environmental parameters.

2.4.2 Activation Tagging

In a genome, some genes are redundant and some are essential for survival. A simple insertional mutagenesis in such cases will not yield a discernible mutant phenotype. Different kinds of specialized vectors have been devised to trap even such genes by a gain of function approach. Activation tagging, i.e. T-DNA tagging using a reporter gene driven by a minimal promoter was first reported by Hayashi and co workers using a vector pPCVICEn4HPT (Hayashi *et al.*, 1992). Although this paper has since been retracted (Balter, 1999) the basic technique developed by these workers still remains valid. Activation tagging is gaining importance these days as can be seen by the number of genes that have been cloned using this method (Kardailsky *et al.*, 1999; Weigel *et al.*, 2000; Huang *et al.*, 2001 and van der Graff *et al.*, 2000). Different activation tagging vectors have now been designed for example the pMON 29963 (Huang *et al.*, 2001), pSKI074 and pSKI015 (Weigel *et al.*, 2000) conferring resistance to kanamycin and to Basta herbicide respectively. In these vectors four copies of CaMV35S enhancer elements were put within and close to the right border of the T-DNA. Loss of function as well as gain of function mutations can be obtained since insertion of enhancer elements lead to transcriptional activation of nearby genes resulting in a dominant gain of function mutation. More recently 45000 T-DNA insertion lines were generated using the actvation tag vector pSKI015 in *Arabidopsis*. Out of the

1194 TDNA insertion sites 1010 were found to be in or close to a predicted gene and nearly half of these genes were those with unknown functions (Qin *et al* 2003). T-DNA insertional mutagenesis is also being established in rice. A new T-DNA vector, pGA2715 for activation tagging and promoter trapping in rice has been developed (Jeong *et al.*, 2002).

2.4.3 Promoter trapping using T-DNA binary vectors

A promoter-trap vector contains a promoter-less reporter gene fused to the T-DNA right border in a binary vector. A random insertion of the T-DNA containing the promoter less reporter gene within a gene might result in a transcriptional or translational fusion. The expression pattern of the reporter gene thus would reflect the expression profile of the tagged gene, since both are under the control of the same promoter. The *gus (uid A)* reporter gene is the most often used reporter gene system in plants, because of the absence of endogenous β-glucuronidase (GUS) activity in most plants and the opportunity to visualize the presence of the enzyme by histochemical techniques (Jefferson *et al.*, 1987). Enzymatic GUS assays are highly sensitive and the gene product can thereby be accurately and specifically localized.

Since reporter gene expression is dominant and can be monitored in heterozygote plants, promoter trap system is useful for studying the patterns of most genes including essential genes, which otherwise cause lethal mutations. This system is convenient for observing mutant phenotypes because reporter gene activation indicates the location, condition and time of expression of the disrupted gene. In *Arabidopsis*

activation of reporter genes can be as high as 30 % (Sundaresan *et al.*, 1995). The insertion of T-DNA into the tagged gene also facilitates its cloning. Lindsey and co workers have demonstrated the potential of promoter trapping in tobacco, *Arabidopsis* and potato and the value of the approach as a complement to standard T-DNA insertional mutagenesis by analyzing in detail the expression of the functional promoter tag in diverse cell types (Lindsey *et al.*, 1993). Promoter trap lines have proved to be invaluable for the creation of marker lines (Kertbundit *et al.,* 1991; Topping and Lindsey, 1997) and for the isolation of regulatory sequences eventually leading to the isolation of specifically expressed genes (e.g. Topping *et al.,* 1994 and Wei *et al.,* 1997). Promoter trap strategy has also been exploited for the isolation of environmental and hormonal stress responsive regulatory sequences or genes (Lindsey *et al.,* 1993; Mandal *et al.,* 1995). This approach is particularly relevant for identification and cloning of genes (and their regulatory sequences) expressed in tissues that are difficult to analyze, by traditional methods relying on RNA extraction. This fact is demonstrated by the cloning of regulatory sequences driving reporter gene expression in nematode feeding structures (Barthels *et al.,* 1997) or the identification of molecular markers for embryogenesis (Topping *et al.,* 1994; Topping and Lindsey, 1997). A simple strategy based on promoter trapping for isolation of different promoters for driving different selectable marker genes has been designed in *Brassica napus* (Bade *et al.*, 2003). In order to identify genes expressing in dividing plant cells, a system of promoter trap insertional mutagenesis was employed in *Arabidopsis thaliana* (Casson *et al.*, 2002; Lindsey *et al.*, 1993 and 1998 and Topping *et al.*, 1994). This

system, unlike other strategies, does not merely rely upon generation of a mutant phenotype for functional genomics but also facilitates the detection of genes by activation of promoter-less reporter gene by the tagged gene promoter. A population of T-DNA tagged lines using a promoter trap vector, pTLuc has been generated to tag stress responsive genes (Alvarado *et al.*, 2004). Promoter trap markers have been used to investigate mechanisms involved in polar organization (Topping and Lindsey, 1997) and embryogenesis. Promoter trapping has the potential to allow the identification of functionally redundant genes that would not be detectable in a conventional mutational screen (Wei *et al.*, 1997). A T-DNA tagged mutant showing intense GUS expression in young leaves and rapidly growing stem tissues was used to isolate a eukaryotic translation initiation factor *eIF-4AI* (De Greve *et al.,* 2001). A gene identified by promoter trapping, *EXORDIUM*, has been shown to express in proliferating cells and involved in meristem function (Farrar *et al.*, 2003). Table 2 lists a few promoters cloned from plants using promoter-trap vectors.

2.4.4 PROMOTER TRAPPING TO ISOLATE CRYPTIC PROMOTERS

Promoter and enhancer trapping experiments have been performed in plants using a variety of T-DNA vectors that lack promoters or which possess only a minimal 35 S promoter linked to T-DNA border repeats (Goldsbrough and Bevan, 1991 and Herman *et al.,* 1990). It has been generally assumed that activation of promoter less or enhancer less marker genes result from T-DNA insertions within or immediately adjacent to genes. However in many cases, integration of the T-DNA results in expression of the promoterless reporter gene because of the presence of

an otherwise silent regulatory element in the immediate vicinity. Lindsey *et al.* (1993) have referred to these promoters as `pseudo-promoters' and suggested that they may be responsible for activating marker genes in some transgenic lines. Inactive regulatory sequences, buried in the genome, capable of driving gene expression when positioned adjacent to a gene have been described in various organisms and referred to as cryptic promoters. Cryptic promoters have been identified earlier in yeast in the intron of a gene coding for yeast actin (Irniger *et al.,* 1992) and a mammalian melanoma-associated antigen (Takahashi *et al.,* 1991). It is possible that evolutionary processes like insertions and rearrangements could result in presence of cryptic promoters in non-coding parts of plant genomes. There are reports of cryptic promoters in plant genome (Cryptic promoters from plants cloned to date are listed in table 2).

Although cryptic promoters can be cloned by other techniques also, promoter trapping is the most convenient method to tag such promoters in plants. A seed coat-specific promoter in tobacco (Fobert *et al.,* 1994) and a root-specific promoter in *Arabidopsis thaliana* (Mollier *et al.,* 2000) have been isolated from promoter trap lines. Promoter trap lines from *Arabidopsis thaliana* have been used to clone regulatory sequences, possibly cryptic, which mediate guard-cell-specific reporter gene expression (Plesch *et al.*, 2000). It has also been reported that a cryptic promoter from tobacco, tCUP, is able to promote GUS expression in all major organs of transgenic *Arabidopsis* plants (Wu *et al.,* 2001). The cryptic enhancer elements found in the tCUP promoter by a series of deletions (Wu *et al.*, 2003) are found to be similar to elements found in promoters of

plant genes. More recently it has also been discovered that the portion of the tCUP sequence containing a variety of cryptic gene regulatory elements is related to a new family of moderately repetitive sequences the RENT (repetitive element from *Nicotiana tabacum*) family. The tCUP promoter has been well characterized and has been developed into a proprietary gene expression system at the Southern Crop Protection and Food Research Center, Canada.

Table 2. Promoters cloned using promoter trap vectors

S. No.	Promoter	Specificity of expression	Source species	Reference
1	At EM	Embryo	*Arabidopsis*	Topping *et al.* (1994)
2.	**Cryptic**	**Seedcoat specific**	***Nicotiana tabacum***	**Fobert *et al.* (1994)**
3	HVT1	Tapetum and vascular tissue	*Arabidopsis*	Wei *et al.* (1997)
4	**Cryptic**	**Callus and roots**	*Arabidopsis*	**Okresz *et al.* (1998)**
5	Pyk20	Nematode feeding structure	*Arabidopsis*	Puzio *et al.* (1999)
6	**tCUP(Cryptic)**	**Constitutive**	***Nicotiana* tabacum**	**Foster *et al.* (1999)**
7	**Cryptic**	**Guard Cell**	***Arabidopsis***	**Plesch *et al.* (2000)**
8	**Cryptic**	**Roots**	***Arabidopsis***	**Mollier *et al.* (2000)**
9	Lj Cbp 1	Roots	*Lotus Japonicus*	Webb *et. al.* (2000)
10	eIF-4A1	Growing tissues, young leaves	*Arabidopsis*	De Greve *et al.* (2001)
11	Exordium	Proliferating Cells	*Arabidopsis*	Farrar *et al.* (2003)

2.5 IMPORTANCE OF TISSUE-SPECIFIC PROMOTERS

Transgenics are being created in different crop plants for a variety of purposes. In a majority of the transformation experiments and in transgenic plants, the promoters used are derivatives of CaMV35S. In monocots, the promoter used has been either a derivative of actin or ubiquitin promoter. The promoters used so far, to a very large extent drive constitutive expression of the transgene. In order to ensure the expression of the introduced gene in a given tissue at a specific time it is essential to identify and isolate a repertoire of DNA elements, which once fused to a coding sequence regulate the expression of this gene in a highly controlled manner. Tissue-specific promoters control gene expression in a tissue dependent manner and according to the developmental stage of the plant. Transgenes driven by these kinds of promoters will only be expressed in the tissue where their expression is desired. The identification of novel genes and a variety of regulatory elements of DNA required for controlled expression of the introduced gene assumes great importance. Many more regulatory elements that impart a tissue-specific, stage-specific and/or environmental-stimuli-specific expression to the transgene must be identified and cloned. Tissue-specific expression is the result of interaction of several levels of gene regulation. Several patents have been filed for different tissue-specific promoters, which includes root, fruit and seed specific promoters. An exhaustive list of patents filed related to tissue-specific promoters and also other inducible promoters are listed in a document of CAMBIA, Australia (Roa-Rodriguez, 2003). The study of DNA sequences regulating gene expression in specific ways is crucial for

understanding the morphogenetic pathways of plant differentiation and for plant genetic engineering. Differential gene expression includes expression restricted to certain tissues or particular stages of plant development, as well as expression in response to specific environmental stimuli. Different tissue-specific promoters have been thoroughly studied. Step by step deletion of different regions of these promoters reveal the specific elements conferring tissue-specificity or the ability to regulate gene expression at various developmental stages. A 5′ upstream element of a pyrroline 5′ carboxylate reductase gene has been shown to drive tissue specific expression of CaMV35S minimal promoter (Hua *et al.*, 1999) similarly the promoter for S- Adenosyl L- Methionyl Synthetase of *Arabidopsis* has been shown to drive GUS expression in phloem and cortex regions of transgenic poplar (Vander Mijnsbrugge *et al.*, 1996). Fruit-specific promoters have been cloned in tomato (Bird *et al.*, 1988; Deikman and Fischer, 1988 and Blume and Grierson 1997) and in other fruit crops like avocado (Cass *et al.*, 1990). The tissue specific effects of the major seed storage protein in soybean, beta-conglycinin have been studied in transgenic *Arabidopsis* (Ohkama *et al.*, 2002). Deletion analyses have revealed that a 45-bp proximal region containing AACA and GCN4 motif is sufficient to confer endosperm-specific expression of the rice storage protein glutelin gene, GluA-3 (Yoshihara *et al.*, 1996). The combination of the two motifs confers endosperm specificity; functions of the two motifs were dependent on the orientation and/or distance from a G-box element suggesting a synergistic interaction between the factors that recognize those motifs and the G-box element. The promoter of cotton alpha-globulin, an abundant seed storage

protein in cotton has been analyzed and found to be to be highly specific to seeds and its activity increased until embryo maturation in three different plants viz. Cotton, *Arabidopsis* and Tobacco (Sunil Kumar *et al.*, 2002).

Production of transgenic plants benefit from the availability of regulated promoters that determine tissue or temporal specificity of transgene expression. Tissue-specific promoters thus go a long way in conserving plant's overall energy expenditure. Tissue-specific promoters are also useful in developmental genetics where they have been employed to spatially control expression of a toxin for specific genetic ablation for example a diptheria toxin under a root cap specific promoter (Tsugeki and Federoff, 1999) and BARNASE gene under an AtLTP (Lipid Transfer Protein) promoter (Baroux *et al.*, 2001). Plant derived tissue-specific promoters would be an environment friendly alternative to the constitutive CaMV35S, which has been debated to be a hazard to the environment (Ho *et al.*, 1999).

2.6 ROOT DEVELOPMENT

Plant organs are formed from meristems by a regulated program that controls the timing of cell division the orientation of a plane of cell division and the extent of cell expansion (Steeves and Sussex, 1989). Roots support the plant, synthesize hormones, acquire water and are the sites of interaction with soil bacteria. Root development is a continuous process in which different cell types arise from a population of progenitor cells (Esau, 1977). Although it is difficult to study an organ that usually grows underground, there are some good reports of genetic analyses of root development (Benfey *et al.,* 1993; Okada and Shimura, 1990 and

Scheifelbein and Benfey, 1991), particularly in *Arabidopsis* whose small size and a very simple root architecture makes it possible to screen large numbers of roots for abnormal development. Even at a low magnification, regions or zones of root development are readily apparent. The meristematic zone is the region in which the earliest detectable progenitors or initials of the differentiated cells are located. In the elongation zone, cell division and cell elongation take place in a precisely coordinated fashion. This zone is characterized by the presence of smaller cells with dense cytoplasm. Towards the tip of the elongation zone, cells begin to acquire their final differentiated attributes. The specialization zone is the name of the region in which the cells begin to become fully differentiated (Steeves and Sussex, 1989). Root morphogenesis mutants have been characterized in *Arabidopsis* using monoclonal antibodies to membrane and cell wall components and also by creating transgenic tissue-specific marker lines (Benfey *et al.*, 1993). Root developmental mutants have defects in genes that govern organ formation, meristem activity, cell differentiation and response to environmental conditions. A gene responsible for root meristem formation (Willemsen *et al.*, 1998) and another gene for cell expansion (Schindelman *et al.*, 2001) in roots of *Arabidopsis* have been thoroughly characterized. Different genes and signaling pathways involved in root differentiation and patterning have been thoroughly reviewed (Casson and Lindsey, 2003). A high-resolution gene expression profile of different cell types in *Arabidopsis* root has also been generated (Birnbaum *et al.*, 2003). Thus root has now become a model organ for developmental studies due to its amenability to anatomical and genetic analyses and the

availability of several transgenic marker lines. Studies of genes and promoters related to root development have an immense scope in plant genetic engineering.

2.7 IMPORTANCE OF ROOT-SPECIFIC PROMOTERS

Several root specific promoters have been identified and characterized (Patent Application by Pioneer Hybrid: US 2001047525 EP 1 248 850). These include promoters that contain root specific elements and also promoters with root-preferred elements that enhance or suppress the expression of the associated gene in roots. Root-specific promoters are very useful in devising newer strategies for genetic engineering. The importance of root-specific promoters is covered in the following sections.

2.7.1 Root specific promoters in nutrient acquisition

Roots absorb nutrients from the soil for plant growth and development. The structure of a root system is so designed as to enable it to forage for mineral nutrients. Nitrate is the major source of mineral nitrogen for plants. Plants respond to nitrate levels in soil in different ways. The plant nitrate uptake system must be robust due to fluctuating nitrate levels in soil. Nitrate uptake in plants has been a subject of study since long (Crawford, 1995) and different genes involved and their regulation has also been reviewed (Orsel *et al.*, 2002). An *Arabidopsis* MADS box gene that induces proliferation of lateral roots in response to low levels of nitrate has been cloned (Zhang and Forde, 1998). A T-DNA line in *Arabidopsis thaliana* where the complete AtNRT2.1 gene and the 3' region of AtNRT2.2 gene (genes involved in nitrate transport) has been knocked out has been identified (Filleur *et al.*, 2001). This mutant is shown to be impaired in

nitrate transport. The promoter of nitrate transporter gene AtNRT2.1 from *Arabidopsis* directs reporter gene expression in epidermal, endodermal and cortical cells of mature root (Nazoa *et al.*, 2003). The cryptic root specific promoter (Mollier *et al.*, 2000), identified by screening a population of promoter trap lines could be actually regulating a nitrate inducible gene (*NIA*1) situated upstream of the promoter, studies are on in order to confirm this hypothesis. The second major nutrient is phosphorous; phosphate uptake activity increases in response to phosphate starvation. Two root specific genes MtPT1 and MtPT2 of the model legume *Medicago truncatula* have been shown to be up regulated in response to phosphate starvation (Liu *et al.*, 1998). The promoters of MtPT1 and MtPT2 have been cloned and shown to express specifically in roots of *Arabidopsis* and *Medicago truncatula* (Xiao *et al.*, 2003). Acid phosphatases are also known to increase during phosphate starvation (Ascencio 1997). The role of an acid phosphatase in Indian mustard (*Brassica juncea*) has been investigated and the levels of two acid phosphatases were found to increase in response to phosphate starvation. The promoter of an acid phosphatase homologue in *Arabidopsis* has been fused to GUS reporter gene. This promoter drives expression of GUS in roots especially at lateral root meristems under phosphate starvation. The expression diminished with addition of phosphate (Haran *et al.*, 2000). Another important nutrient is potassium. Potassium channels encoded by gene AKT1, identified by a reverse genetics approach (Hirsch *et al.*, 1998), mediates potassium uptake in plants. The AKT1 promoter has also been shown to direct expression of GUS in mature roots of transgenic *Arabidopsis* plants

(Lagarde *et al.*, 1996). Other sets of genes, AtKT1 and AtKT2 (Qinter and Blatt, 1997) and AtKUP family (Kim *et al.*, 1998) also mediate potassium transport in *Arabidopsis*. Most of these genes involved in nutrient transport are mainly expressed in roots indicating the importance of roots in nutrient absorption. The promoters of these genes could have an important role in producing better-equipped transgenic crop plants in nutrient deficient conditions.

2.7.2 Root Specific promoters in controlling plant parasitic nematodes

A promoter that provides maximum expression in nematode feeding sites is a basic requirement for engineering plant resistance against parasitic nematodes. Constitutive promoters allow expression of pre-formed defenses in which the effector (nematicidal) protein is present continuously throughout the plant. Biotechnology offers several sustainable solutions to control plant parasitic nematodes. A thorough review (Atkinson *et al.*, 2003) dwells on several aspects of nematode control using molecular biological tools. A promoter activated in the galls of root knot nematode in tomato has been identified (Escobar *et al.*, 1999). The CaMV35S promoter from cauliflower mosaic virus is a widely used example of a constitutive promoter and has been used to demonstrate the efficacy of proteinase inhibitor defense. However, it has already been shown that this promoter is down-regulated in nematode feeding cells (Goodijn *et al.*, 1993). The wun1 promoter shows some promise as it responds to nematode infection, however its activity is lost once the nematode enters the phase of syncitial induction (Hansen *et al.*, 1996). Such a promoter could be put to use in control of invading nematodes but not in established parasites like the

sedentary nematodes. Promoters of viral origin that are effective in expressing linked protein sequences in nematode feeding sites have been identified by Dr. Carmen Fenoll and colleagues at the Department of Biology of Universidad Autónoma de Madrid (www.plantbioscience.com/techs/pdf/16.pdf). The promoters are induced at early stages of nematode infection specifically at nematode feeding sites in various plant species as evidenced by Promoter:GUS fusion studies in *Arabidopsis*. The promoter for a phosphoribosyl formyl glycinamidine (FGAM) synthase (EC 6.3.5.3) from soybean directs the expression of reporter gene in syncytia in *Arabidopsis* plants challenged with *Heterodera schachtii* (Vagchippawala *et al.*, 2004). Thus root specific promoters could be used to spatially limit the expression patterns of the effectors in roots in areas where the nematodes attack.

2.7.3 Root-specific promoters in Storage Organs

For targeted engineering of plants, tissue specific promoters are very useful. For example in plants like cassava, carrot etc., with storage roots, a root specific promoter could be used to improve root composition and quality by genetic engineering. In cassava two promoters have been recently cloned and been reported to be related to vascular expression and storage root formation (Zhang *et al.*, 2003). Most storage roots are rich in starch but are poor in protein and other micronutrients. Moreover the storage roots, once harvested are subjected to rapid post-harvest deterioration. The expression level of CaMV35S promoter has been reported to be lower in storage root in Cassava (Zhang *et al.*, 2003) thus making it imperative to uncover root-specific promoters. It has been

reported that prosystemin gene from tomato under the control of CaMV35S promoter is able to dramatically increase storage protein levels in potato tubers (Narvaez-Vasquez and Ryan, 2002). Sporamin (tuberous root storage protein genes of sweet potato) is usually not expressed in organs other than roots in field-grown sweet potato, however the stems of sweet potato were shown to accumulate large amounts of this protein when grown *in vitro*. Studies by Hattori *et al.* (1990) have revealed that a portion between -90 and –390 in the sporamin promoter are needed for the expression of sporamin in stems of tobacco.

2.7.4 Root-specific promoters in symbiosis

Leguminous plants and *Rhizobia* share a unique non-pathogenic relationship. These bacteria dwell in specialized structures called nodules on the roots of these plants helping the plant by providing nitrogen nutrition. The nuances of this wonderful relationship called symbiosis have been reviewed thoroughly (Hirsch *et al.*, 2001. Several other grass species like rice, maize and sugarcane (actinorhizal plants) associate with nitrogen-fixing actinomycetes. *Actinomycetes* however do not form nodules. The establishment of symbiosis requires continual signal exchanges between the bacterium and the plant. Signaling molecules from the plant, flavonoids interact with the bacterial *nod D* gene and help in assembling a specific factor, which is recognized by a particular legume. Many bacterial signal transduction genes for the symbiosis, such as *nod* and *exo* genes have been identified. Many plant genes have also been implicated at different stages of nodulation response. The hunt for a Nod factor receptor is on for quite a long time. Several studies towards this objective are under way

(Cullimore *et al.*, 2001). A gene for Ca^{2+}/calmodulin-dependent protein kinase involved in nodule formation in *Medicago truncatula* has been cloned (Mitra *et al.*, 2004). The mutant *dmi3* fails to induce nodule production although other changes in response to Nod factors are observed. Thus it has been hypothesized that this gene could trigger other downstream genes. Root specific promoters could be extremely useful in studying the expression patterns of such genes and hence help in determining the whole gamut of plant genes involved in the nodulation factor response. The promoter of soybean early nodulin gene *enod2B* is induced by rhizobial factors in transgenic rice specifically in roots. The expression also was regulated by nitrogen status signifying a signal transduction response (Yanzhang *et al.*, 2004). A promoter-tagging program in the legume *Lotus japonicus* has been initiated to identify plant genes involved in the nitrogen-fixing symbiosis between legumes and *Rhizobia*. Four out of seven tagged plant lines expressed GUS specifically in roots and/or nodules only after inoculation with nodule-forming *Mesorhizobium loti*. In one of these lines the T-DNA insertion was located upstream of a coding sequence similar to calcium binding proteins (Webb *et al.*, 2000). A trigger for the promoter of the *Vicia faba* L. leghemoglobin gene VfLb29 is conserved in leguminous and non-leguminous plants. In both the species they are activated specifically in infected cells of root nodules and in the arbuscule-containing cells of mycorrhizal roots from different legume and nonlegume plants (Tobacco) (Vieweg *et al.*, 2004). An analysis of expression pattern of MtSucS1, a gene encoding sucrose synthase during endosymbiosis, has been carried out in the model legume

Medicago truncatula by fusing it to a *gus*-int reporter gene. The MtSucS1 promoter expressed very strongly in the cortical cells of roots infected with endomycorrhizae specifically in highly colonized root sections. GUS-expression was also detected in the surrounding cortical cells, irrespective of direct contact with the fungus (Hohnjec *et al.*, 2003).

2.7.5 Phytoremediation and Molecular Farming

Several studies in recent times (Gleba *et al.*, 1999 and Drake *et al.*, 2003) have focused on use of plants in alternative agriculture viz. beyond food and fiber. This is beginning to change plant biology in a big way. The use of plants to extract pollutants from soil, called phytoremediation, has been used to cleanse soil of heavy metals (Clemens *et al.*, 2002 and Khan, 2000). A natural ability to hyperaccumulate metals in plants grown in soils containing heavy metals is reported (Gleba *et al.*, 1999). It is also known that plants having higher transpiration rates have a better ability to extract pollutants from soil. Plants like tobacco, maize, sunflower and mustard are used to accumulate metals based on their adsorption capability (Meagher, 2000). Organic pollutants can also be extracted from soil by plants. Tree species like poplar, willow as also alfalfa and different grass species have been used to mineralize organic pollutants. Four different genotypes of *Sorghum bicolor*, with differing root characteristics planted in soil contaminated with crude oil have been shown to dissipate significant amounts of crude from the contaminated soil (Banks *et al.*, 2003). The rhizosphere, which contains various nutrient-rich root exudates, is the region of high microbial activity. The role of different plants in remediation of various environmental pollutants has been reviewed and different studies

on the effect of rhizosphere microbes on plants and vice versa has been documented (Kuiper *et al.*, 2004). A plant-bacterial bioremediation system has been developed in *Astragalus sinicus*, a green manure, by expressing a human metallothionein (MTL4) under nifH and nodB promoters, in nodule forming *Mesorhizobium huakuii*. It was found that symbionts showed increased heavy metal accumulation compared to non-symbionts (Sriprang *et al.*, 2002). The process of competitive root colonization has been studied to unravel the molecular mechanisms and reviewed (Lugtenberg *et al.*, 2001). Discovery of new genes specifically expressed in roots, whose proteins are secreted into the rhizosphere and their interaction with bacteria would help in unraveling the mechanism of phytoremediation. A gene specific to maize root cap and secreted in the root mucilage, ZmGRP4 has been cloned (Matsuyama *et al.*, 1999). Root specific promoters fused to reporter genes could be used to study the interaction of root colonizing bacteria and their role in dissipating pollutants. The other process that plants could be put to use is in the production of recombinant proteins and valuable organic materials. Giant networks of roots in soil act as huge pumps that extract and concentrate essential nutrients from soil. Roots also secrete several substances in order to take up nutrients and also to protect the plant from numerous plant pathogens. The roots have evolved sophisticated mechanisms to secrete different substances. Borisjuk *et al.* (1999) have demonstrated that root secretions can be successfully exploited for the production of recombinant proteins in a process called "rhizosecretion". It was also demonstrated that the recombinant proteins was found to be higher in the exudates than in the roots themselves.

Rhizosecretion can be used as a phytomanufacturing technique and can be designed to produce valuable organic compounds from roots. The process of purification of proteins from the rhizosphere would be much cheaper than production *in vitro* using transgenic plants or aseptically grown plant organs. Since various root exudates are known to have moderate to high activity against a variety of microbes, rhizosecretion could be used to produce pharmaceutically valuable compounds

A variety of root-specific promoters are required for different applications. A root specific promoter, which is wound inducible, would be very useful in nematode control. A promoter specifically expressing in vascular tissues of a root makes it an ideal candidate for use in engineering storage roots. Nodule and nitrogen fixation related promoters could also prove to be very useful in improved nitrogen fixation in plants. Thus, depending upon the nature of root specific expression of the respective promoters, they could be put to several biotechnological applications, like pest and disease control, better nutrient absorption and storage, phytoremediation and molecular farming.

MATERIALS AND METHODS

Seeds of *Arabidopsis thaliana* ecotype Columbia were initially obtained from Prof. J P Khurana, Department of Plant Molecular Biology, Delhi University, South Campus. Soilrite was procured from Keltech Energies Ltd., (Perlite Division) Bangalore and consisted of peat moss, perlite and vermiculite in 1:1:1 proportion. *Arabidopsis* plants were grown in growth chambers and greenhouses of the National Phytotron Facility, IARI, Delhi. Sequencing was carried out at the sequencing facility, Delhi University, South Campus, New Delhi.

3.1 CHEMICALS

Vitamins, hormones, antibiotics and other molecular biology grade chemicals like Fe-EDTA, CTAB (cetyl trimethyl ammonium bromide) were obtained from Sigma Chemical Company (U.S.A.). X-Gluc (5 bromo-4 chloro -3 indolyl-2-D glucuronide) was purchased from Biosynth. Ethanol was procured from Bengal Chemicals, Kolkota and Merck. Chemicals for preparation of MS medium, nutrient solution and other purposes were procured from Hi Media Laboratories, Mumbai, India.

3.2 KITS AND ENZYMES

Plasmid Midiprep kit, QIAquick PCR purification kit and QIAquick gel extraction kit were purchased from Qiagen Inc. Valencia, USA. Restriction enzymes were obtained from Promega Corporation, Madison, USA, Bangalore Genei, Bangalore and MBI Fermentas Lithuania. Other

modifying enzymes like Klenow and T4 DNA polymerase were procured from MBI Fermentas, Lithuania. *Taq* polymerase was purchased from Bangalore Genei, Bangalore, India.

3.3 BACTERIAL GROWTH MEDIA AND INOCULATION

The fresh standard liquid medium of Luria Broth (LB) Broth was used for *E. coli* growth in selective liquid cultures. Luria agar (LA) was used for selective plate cultures. LB medium was used in techniques like preparation of competent cells, transformation of *E. coli*, plasmid isolation and also for preparing frozen stocks at -70^0C. The ingredients of LB media are given in A.1. The cultures were first streaked on a freshly prepared agar plate containing the appropriate antibiotic such that isolated single resistant colonies were obtained, which were further inoculated for plasmid isolation and confirmation. Fresh glycerol stocks were again prepared from such cultures after successful confirmation by restriction analysis. For *Agrobacterium* too, LB medium was used when used for floral dip otherwise Yeast Extract Mannitol medium (YEM) (A.1.3) was preferred.

3.3.1 Bacterial strains and plasmids

The strains of *E. coli*, *Agrobacterium tumefaciens* and plasmids used in the present study and relevant characteristics are listed in Table 3.

Table 3. Bacteria and plasmids and their relevant characteristics

Sl. No.	Strain	Antibiotic marker	Characteristics	Reference
1.	DH5α	Nalidixic acid	Recombination deficient *E. coli* strain	Hanahan, 1983
2.	pBS	Ampicillin	Col E1 Origin of replication	Stratagene
3	pRN1	Kanamycin	T-DNA based promoter trap vector carrying *gus*-int gene without promoter and Basta as plant selectable marker	Resminath, 2003
4	pCS1	Kanamycin & Ampicillin	Characteristics of pRN1 and ColE1 origin of replication	Present Study
5.	pBI101	Kanamycin	Binary T-DNA vector with *nptII* as plant selectable markers. A gus gene without promoter fused to nos terminators	Jefferson *et al.*, 1987
6.	GV 3101	Rifampicin & Gentamycin	*Agrobacterium* strain derivative of C58, used for *in planta* transformation	Koncz and Schell, 1986

3.3.2 Preparation of E. coli Competent cells

Competent cells were prepared according to the standard $CaCl_2$ protocol (Sambrook *et al.*, 1989). A single *E. coli* (DH5α) colony was picked from an overnight grown LA plate and inoculated in 5ml LB medium and incubated overnight at 37^0 C with shaking at 200 rpm. One ml of this culture was used to inoculate a 50 ml LB medium in a 250 ml flask containing nalidixic acid (15 mg l^{-1}) and grown at 37^0 C for 3-4 hrs with vigorous shaking (200 rpm) till OD_{600} reached 0.5-0.6. All the steps after this were carried out at 4^0 C. The culture was centrifuged in an aseptic centrifuge tube at 6000 rpm for 5 min. To the pellet, 10 ml of 0.1M $CaCl_2$ was added,

shaken well to dispense the pellet and incubated in ice for half an hour. The tube was then centrifuged again at 6000 rpm for 10 min. The pellet was dispensed in 1 ml of 0.1 M $CaCl_2$, chilled in ice overnight and then 100 µl aliquoted in sterile micro centrifuge tubes. The micro centrifuge tubes containing the cells were snap chilled in liquid nitrogen and stored at -70^0 C till further use.

3.3.3 Transformation of Competent E. coli

Roughly 1 µg of DNA was added to a microcentrifuge tube containing freshly thawed competent cells and kept on ice for 10 min. Then the cells were transferred to a water bath maintained at 37^0 C for 90 secs and immediately chilled in an ice bath. About 800µl of LB broth was added to the micro centrifuge tube and shaken at 200 rpm in 37^0 C for 1 hr, it was then centrifuged and the pellet dispensed in 100 µl of LB and plated in LA plates containing appropriate antibiotics.

3.3.4 Construction of pCSI Promoter trap vector

1µg of pRN I plasmid, which was already available in our laboratory, and 1µg of pBS (Stratagene) was digested with *Hin*dIII. A 20 µl digestion reaction was set up as follows:

Plasmid DNA: 2µl (1µg)

Enzyme Buffer: 2µl

Restriction Enzyme 0.5 µl (5 units)

Sterile Distilled water to 20 µl

The two digestion mixtures were purified. 5 µl of digested pBS and 2 µl of digested pRNI were kept for ligation using T4 DNA ligase from MBI

Fermentas. The manufacturer's instructions were followed. The 10 µl ligation reaction was set as follows:

Digested DNA: 5 + 2 µl

Ligase buffer: 1µl

T4DNA Ligase: 1µl

Sterile Distilled water: to 10µl

The ligation was carried out overnight at 8^0 C. The ligation product was used to directly transform competent *E. coli* cells and plated on LA plates containing required antibiotics. The plasmids isolated from colonies obtained were analysed by restriction digestion.

3.3.5 Isolation and purification of plasmid DNA

Plasmid DNA was isolated by the alkaline lysis method (Birnboim and Doly, 1979). 5 ml of LB in a vial was inoculated with a single colony of *E. coli* carrying the plasmid and grown overnight at 37^0C. The culture was centrifuged in a micro centrifuge and the cell pellet was properly resuspended in 150µl buffer P1. The tube was then incubated for 10 min on ice. 150 µl of buffer P2 was added and the contents were gently but thoroughly mixed by inverting the tube several times and incubated at room temperature for 5 min. To this, 150 µl buffer P3 was added and the contents were again mixed. After 10 min the tube was spun in a micro centrifuge at 12000 rpm for 15 min. The supernatant carrying the plasmid was carefully transferred to a new tube and 0.6 volumes of isopropanol was added and mixed. The contents of the tube were centrifuged again at 15000 rpm for 20 min. The pellet was washed with 500 µl 70% ethanol. After a spin at 12000 rpm for 10 min, the pellet was air dried and dissolved

After a spin at 12000 rpm for 10 min, the pellet was air dried and dissolved in an appropriate quantity (40 µl) of TE. For a diagnostic mini prep this preparation was enough, however in order to use plasmids for cloning or ligation the dissolved plasmid was further purified. The plasmid was made up to 500 µl with distilled water to which, an equal quantity of a mixture of phenol:cholorform:isoamylalcohol (25:24:1) was added, mixed well and centrifuged. The upper aqueous phase was carefully pipetted out in to a fresh tube and extracted with an equal quantity of chloroform twice. The aqueous phase was then taken in a fresh tube and twice the volume of absolute ethanol and 0.1 volume of 3M CH_3COONa added. This was incubated in -20^{0}C overnight and then centrifuged at 12000 rpm for 20 min, the pellet washed with 70% ethanol, air dried and dissolved in an appropriate quantity of TE. (Composition of buffers is given in A.1.4). In case of plasmids like pBI101, which are very low copy number plasmids, QIAGEN plasmid purification kit was used and the protocol for a midi prep of low copy number plasmids given by QIAGEN was followed.

3.3.6 TRANSFORMATION OF AGROBACTERIUM USING FREEZE THAW

3.3.6.1 Preparation of competent cells

200 ml LB was inoculated with 1 ml of an overnight culture of GV3101 *Agrobacterium* and incubated with vigorous shaking at 28^{0}C. Cells were grown to log phase (OD_{550} 0.5-0.8) and pelleted in a bench top centrifuge at 5000 rpm for 10 min at room temperature. The pellet was

washed with 0.1M $CaCl_2$ and the cells resuspended in $1/10^{th}$ original volume of LB and 100 or 150 µl was aliquoted in micro centrifuge tubes.

3.3.6.2 Transformation of Agrobacterium

The competent *Agrobacterium* cells were thawed on ice and 1-3 µl DNA was added to a 150 µl aliquot. The mixture was kept on ice for 5 min and then frozen in liquid N_2 for another 5 min, then the mixture was incubated for 5 more minutes at 37^0C in a water bath. 1 ml of LB was added to each tube, sealed well and the tubes were kept on a shaker (200 rpm) at 28^0C for 6 hrs. The cells were collected by spinning briefly in a micro centrifuge and were spread on two YEM agar plates containing appropriate antibiotics. The plates were incubated for 2 days at 28^0C. 5ml liquid cultures were grown from the colonies and used for carrying out mini-preps, alternatively the colonies were picked using a toothpick, suspended in 10 µl TE and boiled in a water bath for 5 min. This preparation was used as the template DNA for colony PCR with *gus* specific primers. The colonies were streaked on a new plate (master plate) containing the antibiotics and incubated for 2 days at 28^0C.

3.4 PLANT GROWTH MEDIUM

Arabidopsis seeds were germinated on solid Murashige and Skooge (MS) medium (A.2.1). The plants were watered with Hoagland's solution (A.2.2).

3.4.1 Growth of Arabidopsis

3.4.1.1 Seed germination on MS medium

Arabidopsis seeds were surface sterilized by treating them for 2 min in 70 % ethanol followed by treatment with 0.1 % SDS-$HgCl_2$ for 6-7 minutes. The seeds were then thoroughly washed using sterile distilled water with five changes. The surface sterilized seeds were then suspended in 0.1 % agarose solution and spread evenly on Petri plates containing solid MS. These plates were kept for 48 hours at 4^0C to synchronise germination and then incubated in the culture room till the seedlings established themselves, at 22-24^0C at a photoperiod of 16 hrs light and 8 hrs darkness.

3.4.1.2 Planting in soil

Pots (6 cm diameter) were filled to about 0.75 cm from the top with soilrite. The pots were covered with paper and autoclaved at 15 psi, 120^0C. 8-10 such pots were arranged in a tray and saturated with 4-5 l of Hoagland's solution. A thin layer of fine soil was spread over the surface of soilrite. The surface of soil was uniformly sprayed with Hoagland's solution using a spray bottle. Seedlings were gently pulled out of the agar with the help of a blunt forcep, the residual agar removed and then transplanted into the pots containing soilrite (about 3 or 4 plants were carefully placed in a pot). Pots were then covered with Saran wrap secured with rubber bands, so as to keep humidity high. Water was left standing on the tray, the pots were placed in growth chambers maintained at a temperature 20-23^0C and a photoperiod of 16 hr light 8 hr darkness. A relative humidity of 60-75% was maintained and cool white fluorescent tubes, giving 150 $\mu mol/m^2/s$

(1μmol/m^2/s = 53.8 lux) light intensity was also provided. Slits were made on the saran wraps on the 5th day and the wraps were removed by the 7th day. While the plants were young, soil was kept moist by regularly pouring Hoagland's solution in the tray. When the plants established themselves, soil was partially dried out periodically so as to control algal and fungal growth. Supply of water was alternated with that of Hoagland's solution. The plants formed rosettes, inflorescence stalks and flowered generally in 3-4 weeks time and were harvested within 8 weeks. In case of aphid infestation, 0.1% Metasystox™ (Bayer, Germany) was sprayed.

3.4.1.3 Seed harvesting, handling and preservation

Watering was continued till seed set and was discontinued for two weeks prior to harvest. For seed collection, the inflorescences were trained into long inverted butter paper bags. The base of the bags was cut open and the mouth was fastened to the base of the plants. Seeds were harvested by cutting the entire inflorescence at its base after all the siliques became mature and brown. The harvested plant material was allowed to dry for a few days in the bags before threshing. When the inflorescences in the bag appeared to be dry, they were gently hand pressed from outside and seeds allowed to fall to the bottom of the bag. Seeds were sieved to separate them from chaff and then further cleaned by gentle blowing. An additional sieving occasionally was required to clean the seeds completely. The seeds were then air dried for 1-3 weeks at room temperature to reduce the moisture content. Dried seeds were dispensed in micro centrifuge tubes and sealed tightly with PARAFILM™ to prevent dehydration. Every tube

was labeled with relevant information, including date of storage using a waterproof permanent marker and stored in desiccators at 4^0C.

3.5 *IN PLANTA* TRANSFORMATION OF *ARABIDOPSIS*

3.5.1 Preparation of Arabidopsis Plants

For *in planta* transformation of *Arabidopsis,* floral dip protocol (Clough and Bent, 1998) was followed. The plants were placed in pots at a density of 3-4 plants per pot. The first inflorescence shoots were removed as soon as they emerged so as to encourage growth of more shoots. In a week, the plants were ready for transformation.

3.5.2 Preparation of Agrobacterium

Three days prior to plant transformation, a 5 ml LB broth containing appropriate antibiotics was inoculated with *Agrobacterium* carrying the binary vector and incubated at 28^0C with vigorous agitation. After 2 days, 1 ml of this culture was inoculated in 200 ml of LB containing the same antibiotics taken in a 1l flask and incubated again at 28^0C with vigorous agitation (180 rpm) for 24 hrs. The *Agrobacterium* was pelleted by centrifuging at 6000 rpm for 10 min and the cell pellet was resuspended in 400 ml of 5% sucrose. 30 µl of Silwet L-77 was added to this suspension and it was mixed thoroughly.

3.5.3 Floral Dip

The *Agrobacterium* suspension was transferred to a convenient plastic beaker. Whole plants were dipped into the suspension by inverting the pots in the solution such that all the above ground parts were dipped in

the suspension. The plants were left in that position for 1 min. The same suspension was used for ~ 6 pots. After treatment with *Agrobacterium*, the plants were kept for the next 24 hrs under high humid conditions i.e. in a closed box or incubator. After 24 hrs, the leaves were washed with water and the plants returned to their normal growing conditions. The seeds from these plants were collected in about 3-4 weeks.

3.6 SELECTION OF PUTATIVE TRANSFORMANTS USING BASTA

An approximate quantity of seeds was surface sterilized and plated as described above in section 3.4.1.1(@ 100 seeds per plate) in MS medium containing Basta 6 mg l^{-1}. The plates were first incubated in dark at 4^0C for 48 hrs and then transferred to a plant tissue culture room maintained at the standard conditions with adequate light. In about 10 days the resistant plants appeared green and healthy with good roots. The Basta resistant seedlings were then transferred to soil.

3.7 HISTOCHEMICAL GUS-EXPRESSION ANALYSIS

The histochemical GUS assay was carried out as described by (Jefferson *et al.*, 1987). Explants were soaked in GUS assay buffer (A.3) in micro centrifuge tubes and kept in vacuum for 5 min in a desiccator. The tubes were closed and incubated overnight at 37^0C. To remove chlorophylls and pigments the explants were treated with 70% ethanol at 37^0C with two changes (in case of hard tissue temperatures as high as 60^0 C could be used. GUS expression was studied under a Nikon HFX II light stereomicroscope with a fiber optic light source and photographed.

3.8 ISOLATION OF DNA FROM *ARABIDOPSIS*

Genomic DNA of *Arabidopsis* was isolated using CTAB method. The DNA extraction buffer was freshly prepared from the stock solution (A.4) in the final concentration as:

CTAB	2%
Tris- HCl	100mM
Na_2EDTA	20mM
β-Mercapto Ethanol	0.2%
NaCl	1.4M

1-2g of *Arabidopsis* leaf tissue was ground to a fine powder in liquid nitrogen in a pre-chilled mortar. (*For a mini prep 100 mg plant tissue was enough and all the other solutions were scaled down accordingly*). The powder was transferred to centrifuge tubes containing pre-heated extraction buffer (10-15ml at 65°C), mixed well and incubated at 65^0 C for one hour. The tubes, containing the lysed leaf tissue was cooled to room temperature and an equal volume of chloroform:isoamyl alcohol (24:1) added and mixed gently but thoroughly by inverting the tubes several times. The mixture was then centrifuged at 12000 rpm (in a Sorvall RC 5C) for 10 min at room temperature and the supernatant pipetted out carefully in a fresh centrifuge tube, 0.6 volume of isopropanol added and gently mixed. It was incubated at room temperature for half an hour and centrifuged at 12,000 rpm for 10 min for precipitating the DNA. The DNA pellet was rinsed with 70% alcohol, centrifuged at 12,000 rpm for 5-10 min and air-dried. The pellet was resuspended in TE buffer (250-500μl). RNase treatment was given at a rate of 5-10μl of 10mg /ml stock for 500μl DNA sample isolated. The sample was incubated at 37°C for 1 hr, followed by

sequential extractions with phenol:chloroform (1:1), chloroform: isoamyl alcohol (24:1) and chloroform. The final purified aqueous layer was treated with 1/10 volume of sodium acetate (3M) and two volumes of ethanol 100% and incubated at -70°C for half an hour or -20°C overnight. The DNA was pelleted by centrifugation at 10,000 rpm for 10 min. The pellet was washed with 70% ethanol and dried. The pellet was re-dissolved in 100-200µl of TE or SDW, its quality checked on agarose gel and stored at -20°C.

3.9 CONFIRMATION OF T-DNA MUTANTS BY PCR AMPLIFICATION

About 100 ng of DNA of the above preparation was used to check for the presence of *gus* gene in a 25 µl reaction mixture in a 200 µl PCR tube. The reaction was set up as follows:

10 x buffer with 1.5 mM $MgCl_2$	2.5 µl
dNTPs	200 µM
Primers (*gus*-specific)	0.4 µM
Taq DNA polymerase	1.5 units

Annealing temperature of 57^0C was used. The reaction mixtures were kept in an automated thermal cycler (Biometra) and subjected to the following conditions:

94^0C for 4 min

32 cycles of a three-temperature PCR cycle of

94^0C; 30 sec: 57^0C 45 sec and 72^0C 1 minute.

The final incubation at 72^{0}C was for 5 minutes; the reactions tubes were then kept in 4^{0}C. Sequences of primers used are given in A.8.

3.10 CHECKING PCR PRODUCTS

About 10 µl of PCR amplification product was mixed with 2 µl of 6 x loading dye (A.6) and electrophoresed on a 1.0% agarose gel. A 1kb ladder was used as marker. The gel was viewed under an Alpha imager gel documentation system and photographed.

3.11 GEL EXTRACTION

Gel extraction of DNA fragments was carried out using QIA quick gel extraction kit supplied by QIAGEN. The buffers and procedures, recommended by the manufacturer were used. DNA fragment was excised from the agarose gel with a clean sharp scalpel. The gel slice was weighed in a micro centrifuge tube and about 3 volumes of Buffer QG were added to one volume of gel. The tube was incubated at 50^{0}C for 10min until the gel slice dissolved completely. The tube was vortexed every 2-3 min during the incubation. When the gel slice had dissolved completely, 1 gel volume of isopropanol was added to the sample and mixed. A QIA quick spin column was placed in a 2 ml collection tube and the sample (max 800 µl) was applied to the QIA quick column and centrifuged at high speed for 1 min. The flow-through was discarded and the column, placed back into the same collection tube. The column was washed by adding 750 µl of PE buffer. The flow through was discarded and the column was centrifuged for an additional 1 min at 13,000 rpm. The column was placed into a clean 1.5 ml micro centrifuge tube. To elute DNA, 50 µl of elution buffer was added to the center of the QIA quick membrane left at room temperature for 10 min

and then kept over a fresh eppendorf tube and centrifuged for 1 min at maximum speed to collect the eluent in the fresh tube.

3.12 CLONING OF T-DNA FLANKING SEQUENCE

3.12.1 Inverse PCR

Genomic DNA (1µg) isolated from each individual mutant was completely digested with 15 units of *Nde* I enzyme (New England Biolabs). A 30 µl reaction was set according to the manufacturer's instructions and incubated at 37^0 C overnight. 5 µl of the digest was checked on an agarose gel for complete digestion and the remaining 25µl of digested DNA was precipitated with absolute ethanol and re-dissolved in 25 µl water. This digested DNA was used to set up a 15 µl dilute self-ligation reaction. The mixture was incubated overnight at 16^0 C. This self-ligated DNA was used for PCR with TR3 and LB2 primers using Hi Fi Platinum *Taq* polymerase. (Invitrogen). The reaction was set up according to the manufacturer's instructions.

3.12.2 Cloning of PCR products

PCR products were cloned in TA Cloning vector pGEM-T Easy (Promega). The standard procedure given by the manufacturer was followed. pGEM-T Easy vector DNA was briefly spun and a 10 µl ligation reaction set up as given below:

T4DNA ligase buffer:	1 µl
pGEM T-Easy Vector (50 ng)	1 µl
PCR Product	4 µl
T4DNA Ligase	1 µl
Deionized water	to 10 µl

The PCR products were then electrophoresed on an agarose gel as described in section 3.10 and photographed.

3.13 SOUTHERN HYBRIDIZATION

DNA was extracted by CTAB method as described earlier and purified by phenol chloroform extraction and quality checked on agarose gel.

3.13.1 Restriction of genomic DNA and agarose gel electrophoresis

About 6 µg of DNA isolated from leaf tissue of mutants was digested overnight separately with 50 units of *Nde* I and *Hin*d III enzymes at 37^0 C for 6-8 hours. The DNA was electrophoresed on 0.8% agarose gel prepared in 1 X TAE (A.5). 1 kb ladder and lambda *Eco*R I + *Hin*d III digest was used as markers. Electrophoresis was carried out at 40V overnight. The gel was observed under UV and photographed and the marker bands taken note of and the marker portion of the gel cut off.

3.13.2 Blotting of DNA from agarose gel

DNA from the gel was transferred onto a nylon membrane, HyBond N^+ Amersham Pharmacia, UK (Southern, 1975). The gel was treated with excess of 0.25 N HCl for 15 min at room temperature with gentle shaking. After a brief wash with sterile distilled water it was treated in denaturation solution (A.5.3) for 30 minutes with mild shaking. Once again after a brief wash with water it was treated with neutralization solution (A.5.4) Appropriate quantity of 20 X SSC stock (A.5.2) was taken in a glass tray. A glass plate was kept over the tray. Two long wicks, cut from Whatman No. 3 filter paper was laid over the glass plate in a manner that both the ends

are dipped in 20X SSC in the tray. The wick was completely soaked in 20 X SSC. The gel was now placed over the wick and a smooth glass rod rolled over the gel to remove air bubbles trapped underneath. A Hybond N^+ nylon membrane (Amersham) was cut exactly to the size of the gel, soaked in 20X SSC and placed over the gel, air bubbles trapped between the gel and the membrane were removed. Over the membrane, four Whatman No.3 filter papers cut to the same size were placed. A stack of country filter paper also cut to the same size was placed on top of this to a thickness of 5-6 cm. Another glass plate was placed on the top of the filter paper stack over which a 500 g weight was kept. The DNA transfer through capillary action was continued for 16-18 hrs at room temperature. The filter papers were then removed and the portion of the wells on the gel was marked with a soft pencil. The membrane was peeled off the gel and dried at room temperature and stored at 4^0 C till hybridization.

3.13.3 Probe preparation

For probe preparation, PCR was carried out with pCS1 plasmid DNA with *gus* forward and reverse primers and the 1.2 kb amplicon was gel eluted using Qiagen gel elution kit as described before (Section 3.11). The fragment after purification was electrophoresed and the quantity estimated, after comparing with 9 *Hind* III marker. HexaLabel Plus DNA labeling kit (MBI Fermentas) was used to radiolabel the DNA fragment with α^{32}P dCTP. The protocol given by the manufacturer was followed. 150 ng (2 µl) of DNA template, 10 µl Hexanucleotide in 10X reaction buffer and 40 µl deionized water were added to a microcentrifuge tube and vortexed and briefly spun. The tube was then incubated in a boiling water bath for 5-10

min and then chilled on ice. The tube was again spun briefly and to the same tube 3 µl of Mix C and 1 µl of K-^{32}P d CTP and 1µl of klenow fragment was added. The tube was shaken well and then spun down in a micro centrifuge and incubated at 37^{0}C for 10 min. 4 µl of dNTP mix was added and the tube was incubated for 5 more min at 37^{0}C. The reaction was stopped by the addition of 1µl of 0.5 M EDTA, pH 8.0. The labeled DNA was then denatured by keeping in boiling water bath for 5 min and then rapidly cooled and kept at -20^0 C till use in hybridization.

3.13.4 Prehybridization and hybridization

The membrane was incubated with 30 ml of prehybridization buffer (A.5.1) for one hour at 65^0 C in a hybridization bottle kept in a hybridization oven (Amersham) with gentle mixing at 5 rpm. After 2 hours, the denatured probe was added and hybridization continued at 65^{0}C for 16-18 hrs at a slow speed of rotation. The size of *gus* probe was 1.2 kb.

3.13.5 Washing

After completion of hybridization the blot was removed from the bottle and sequentially washed at room temperature for 10 min with 2 X SSC containing 0.1% SDS, for 10 min at 65^{0}C; 1 X SSC with 0.1% SDS for 10 min at 65^{0}C and 0.1 X SSC with 0.1% SDS.

3.13.6 Autoradiography

For autoradiography, the washed membrane was covered with saran wrap and placed inside a lead cassette. An X ray film (Kodak) was carefully placed over the membrane in a dark room. The cassette was kept in -70^{0}C for 3 days to one week depending upon the intensity of the probe.

The film was processed with Kodak developer for 5 min, rinsed in cold water for 1 min and fixed in Kodak fixer for 5 min. The film was washed thoroughly in running water and dried at room temperature.

3.14.1 CLONING OF 5' UPSTREAM REGULATORY SEQUENCES IN pBI 101

Three primers, R1, R2 and R3 were designed in order to clone two sets of sequences upstream of the T-DNA insertion. Three others, AHPB, AHP1 and AHP 0.5 (A.8.3) were designed to clone two sets of the upstream regulatory sequences of the SAHH gene. Primers, R1 and AHPB were provided with a *Bam*H I site at the 5' end while the other primers were provided with a *Sal* 1 site at the 5' end. PCR was carried out with Xt-*Taq* polymerase (Bangalore Genei). The PCR products were checked on an agarose gel and then cloned in UA cloning vector (pDRIVE from QIAGEN). The protocol given by the manufacturer was used. The positive colonies were digested with *Bam* HI and *Sal* I and the released fragments were gel eluted and cloned in the *Bam*H I and *Sal* I site of pBI101. Four such new vectors were derived from pBI.The pBI-derived vectors (pCS2, pCS3, pCS4 and pCS5) were mobilized in to *Agrobacterium* strain GV 3101 by freeze-thaw method and positive colonies checked by colony PCR as explained earlier (Section 3.3.6).

3.15 TRANSFORMATION OF *ARABIDOPSIS*

The *Agrobacterium* strains carrying the new constructs were used to transform *Arabidopsis* by floral dip method as explained earlier (Section 3.5). The seeds collected from transformed *Arabidopsis* plants were selected on MS media containing kanamycin 50 mg l^{-1} as explained earlier (Section 3.6). Histochemical GUS analysis was carried out on the progeny

of the resistant seedlings as explained earlier (Section 3.7) and the GUS expression in roots observed under microscope and photographed.

3.16 *IN SILICO* ANALYSIS

In silico restriction analysis of DNA sequence was carried out using webcutter web based program available in http://rna.lundberg.gu.se/cgi-bin/cutter2/cutter The different promoter regions cloned were analyzed *in silico* using web based programs. For identification of different promoter motifs a program for plant *cis* acting regulatory elements was used available in http://oberon.fvms.ugent.be:8080/PlantCARE/index.html.

RESULTS AND DISCUSSION

In order to isolate genes and promoters using insertional mutagenesis a population of mutants has to be generated. A promoter trap vector pRN1 was available in our laboratory carrying a promoter less *gus*-int reporter gene and *bar* gene (coding for phosphinothricin acetyl transferase protein) as the plant selectable marker (Fig.2). It was decided to construct a new promoter trap vector that contains a plasmid origin of replication in the T-DNA region so that it would be easy to isolate the flanking sequences by plasmid rescue. Towards this end it was decided to introduce a bacterial origin of replication in the T-DNA region of pRN1.

4.1 CONSTRUCTION OF pCS1

In order to introduce a plasmid origin of replication in the promoter trap vector, pRN1 (Fig. 2), the whole pBluescript (pBS) was introduced into the *Hind* III site of pRN1 from pCS1 (Fig. 3). The construction of pCS1 was confirmed by digestion with *Hind* III, which released the complete pBS insert (Fig. 4). The pCS1 vector thus contained two bacterial selectable markers, *Amp* ^{R} and *Kan* ^{R}. The new vector, pCS1 was mobilized into *Agrobacterium* strain GV 3101 and the presence of the plasmid in *Agrobacterium* was confirmed by colony PCR with *gus*-specific primers giving a 1.2 kb amplicon (Fig. 5).

4.2 GENERATION OF A POPULATION OF T-DNA TAGGED LINES

The floral dip transformation system developed by Clough and Bent (1998) has been successfully standardized under our laboratory conditions.

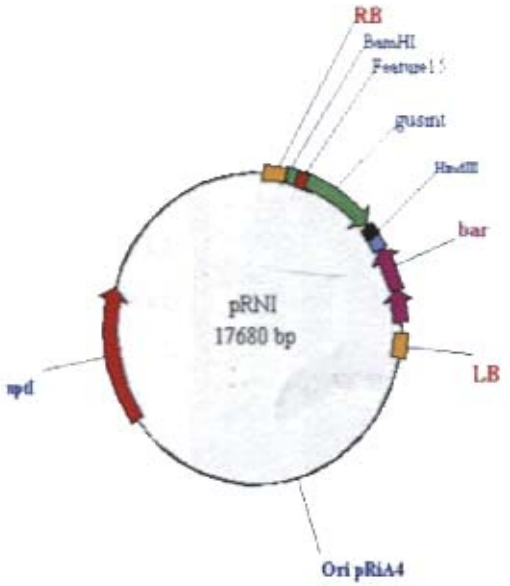


Fig. 2 Map of pRN1 showing the presence of *bar* gene and *gus*-int

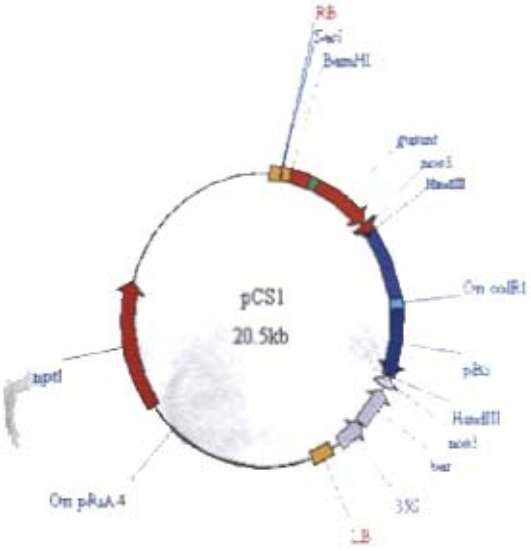
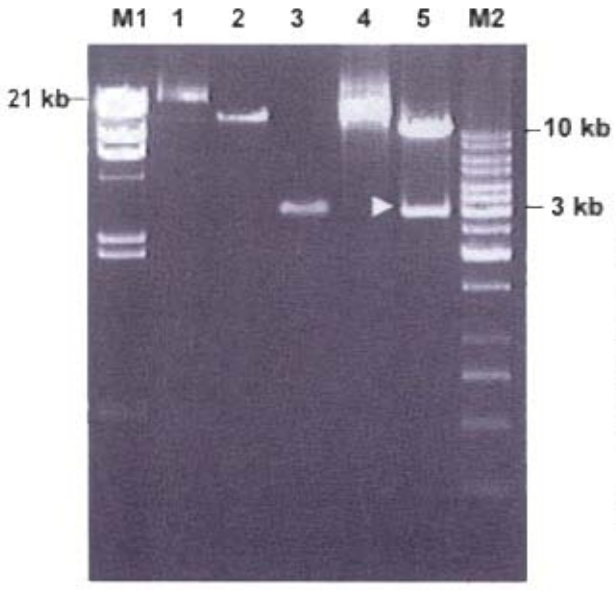


Fig. 3 Map of pCS1 showing the presence of origin of replication



M1:λ *Hind* III digest
1: pRN Uncut
2: pRN- *Hind* III; 17.68 kb
3: pBS- *Hind* III; 2.9 kb
4: pCS- Uncut
5. PCS – *Hind* III; 20.5 kb
White arrow indicates the 3 kb pBS insert
6. M2: 1 kb ladder

Fig. 4 Confirmation of pCS1

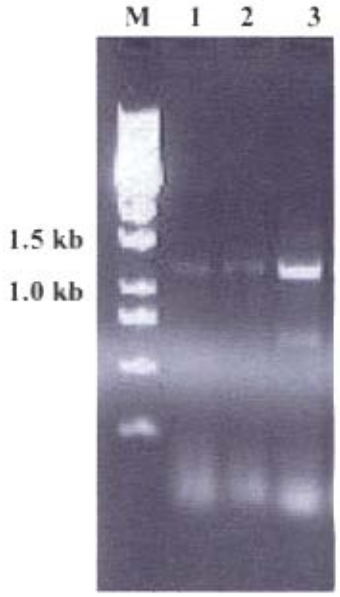


Fig. 5 Colony PCR of *Agrobacterium* containing pCS1 with *gus*-specific primers
M: 1 kb ladder
Lanes 1-3: Positive colonies showing 1.2 kb band

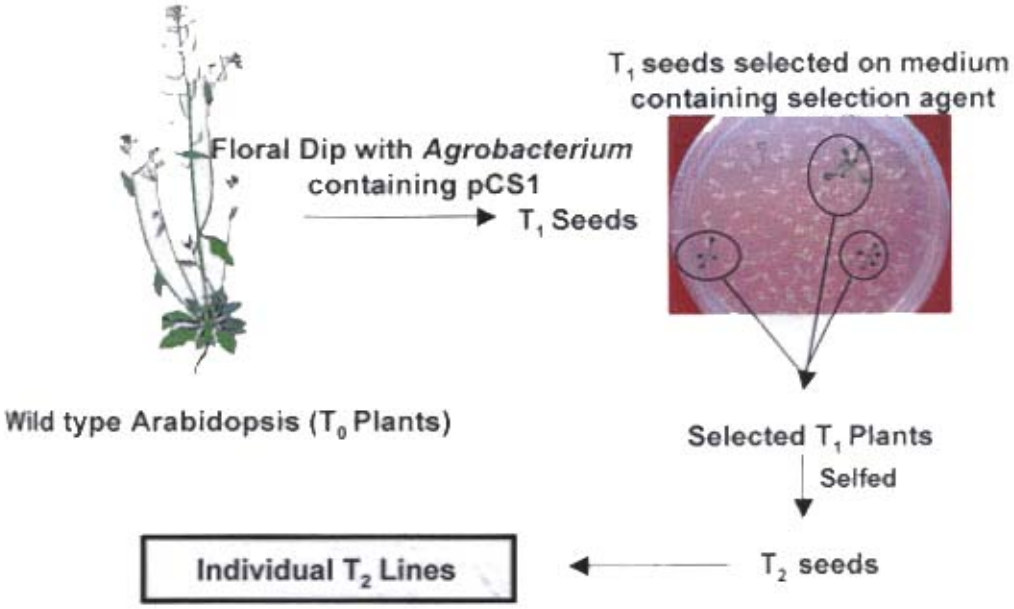


Fig. 6 Scheme depicting the nomenclature used for *Arabidopsis* transformants in this study

In the present study, *Agrobacterium* strain GV3101 carrying pCS 1 vector was used for *in planta* transformation of *Arabidopsis thaliana* ecotype Columbia. The scheme of nomenclature used for *Arabidopsis* transformants is explained in Fig. 6. The wild type plants infiltrated with *Agrobacterium* are referred to as the T_0 plants. The seeds of these infiltrated wild type plants are called T_1 seeds, which are selected on antibiotic to give rise to the T_1 plants. The seeds of the selected T_1 plants (transformants) are referred to as the T_2 seeds, which give rise to the T_2 progeny. Optimum concentration of Basta to be used to select the T_1 transformants was experimentally evaluated by germinating wild type *Arabidopsis* seeds in MS medium containing varying concentration of Basta (3-9 mgl^{-1}). $6mgl^{-1}$ Basta concentration was found suitable for selection of transformants (Figs 7a and b.). The T_1 seeds were selected on Petri plates containing solid MS with 6 mg/l Basta (Phosphinothricin). The resistant seedlings remained green while those that were susceptible turned yellow and died in 15 days. The results obtained with this method were quite comparable to those obtained by sub-irrigating *Arabidopsis* seeds on a sand bed with 10 mg l^{-1} Basta (Bouchez *et al.*, 1993 and Richardson *et al.*, 1998). The T_2 seeds of about 500 Basta resistant T_1 plants (labeled from 1-500) were collected and T_2 lines were generated.

The *in planta* transformation efficiency of pCS1 was found to be 2.0% (Table 4). This is the most common reported transformation frequency of *Arabidopsis* by other workers as Clough and Bent, 1998 have reported a transformation efficiency in the range of 0.5-3% by floral dip

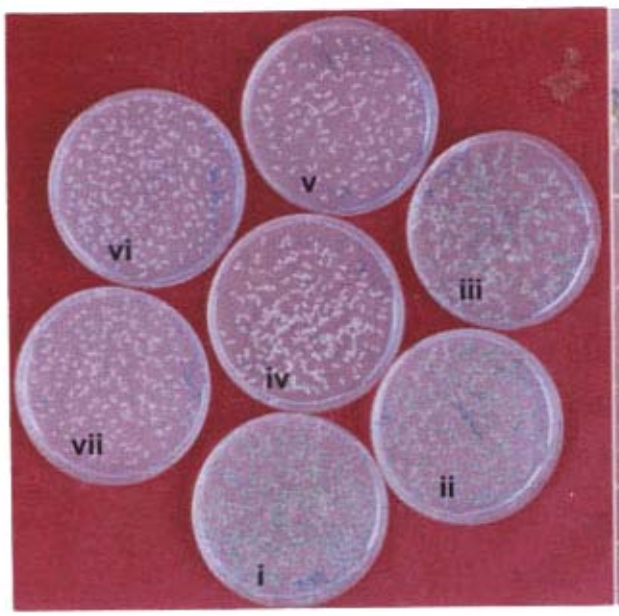


Fig. 7a Wild type seeds germinating on MS medium containing varying concentrations of Basta

i. 2 mg l^{-1}
ii. 4mg l^{-1}
iii. 5mg l^{-1}
iv. 6mg l^{-1}
v. 7mg l^{-1}
vi. 8mg l^{-1}
vii. 10mg l^{-1}

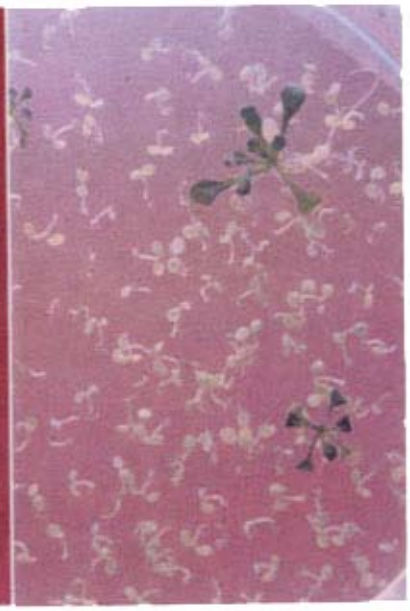


Fig. 7b Selection of transformants at 6mg l^{-1} Basta

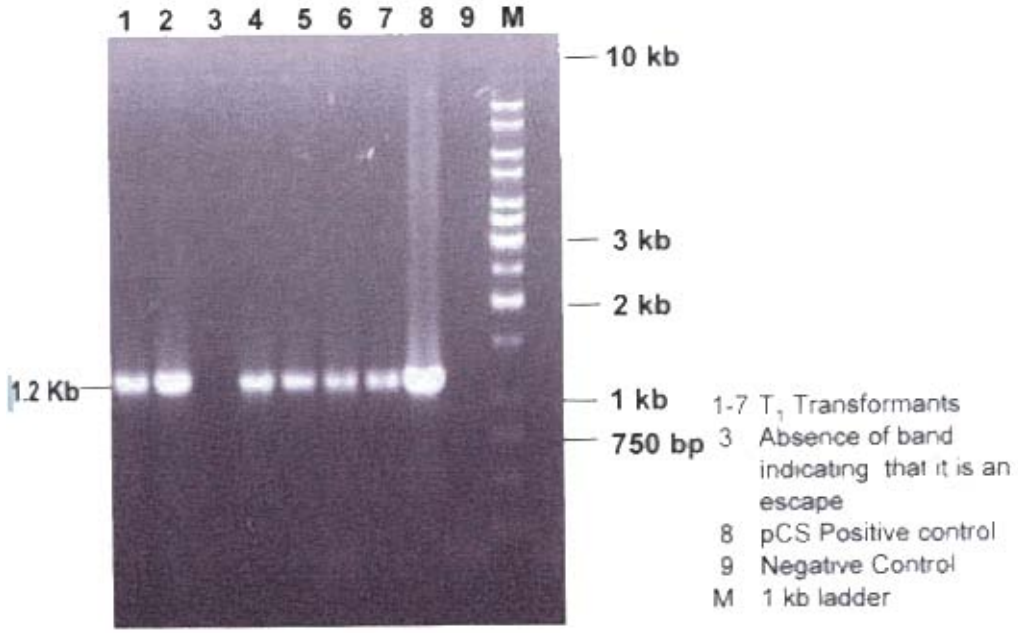


Fig.8 PCR Analysis of T_1 transformants with *gus*-specific probe

method. An earlier work in our laboratory reported that the transformation efficiency was 1 % with pGKB 5 and 2.8% with pRN1 (Resminath, 2003).

Table 4. *In planta* transformation efficiency of *Arabidopsis* with pCS1 (from a sample of plates)

Plate No.	No. of seeds germinated	No. of Transformants	%Transformation
1	130	3	2.3
2	102	2	1.96
3	140	3	2.14
4	100	3	3.00
5	125	1	0.8
6	130	4	3.07
7	120	3	2.5
8	125	4	3.2
9	130	3	2.3
10	120	2	1.67
11	130	2	1.54
12	140	3	2.14
13	125	2	1.48
14	135	3	2.4
15	100	1	1
Total	**1852**	**39**	**2.1**

In the current study, approximately, 30000 T_1 seeds from wild type plants subjected to transformation were screened and 500 primary transformants generated. Of these 500 primary transformants more than 250 were screened for GUS expression in various tissues.

4.3 PCR ANALYSIS

The genomic DNA isolated from T_1 plants was used for PCR with *gus* gene-specific primers. Amplification of a fragment of size, 1.2 kb with *gus*-specific primers suggested the presence of T-DNA in these plants. In a majority of the cases (>90%) the appropriate gene fragment was amplified (Fig. 8), confirming the presence of β glucuronidase gene. The plants in which the amplified products were not obtained were considered as escapes and rejected.

4.4 SEGREGATION ANALYSES AND HISTOCHEMICAL GUS ASSAY

The segregation for Basta resistance was recorded for T_2 plants of all the lines. The segregation ratio in the T_2 provides preliminary information about the number of loci containing the T-DNA in a specific line. Since the resistance marker is dominant, a single insertion in a heterozygous plant will produce progeny that segregates in a 3:1 ratio for $Basta^r$: $Basta^s$, two unlinked inserts will result in a 15:1 $Basta^r$: $Basta^s$, ratio, whereas two linked inserts will fail the test for 15:1 and test positive for the test of linkage. A line for three or four unlinked inserts will segregate in a ratio of 63:1 and 255:1 $Basta^r$: $Basta^s$ respectively. Any failure in agreement to fit the ratios of 63:1 or 255:1 could be interpreted to explain the occurrence of linkage between two or more than two inserts

Table 5 represents the chi–square value for segregation of Basta resistance of a representative sample of T_2 lines. The results show that progeny of most of the transformants have inherited 1-2 copies of T-DNA because most of the lines neither fitted 3:1 nor 15:1 suggesting linked

insertions. Richardson *et al.* (1998), have generated a population of 2165 *Arabidopsis* T- DNA tagged lines and observed that the T-DNA tagged lines contained 1-3 copies of T-DNA as determined by Southern blot analysis using a probe that binds to the junction fragments at the right border of the T-DNA insert.

Table 5. Segregation of Basta Resistance in a selected set of T_2 lines

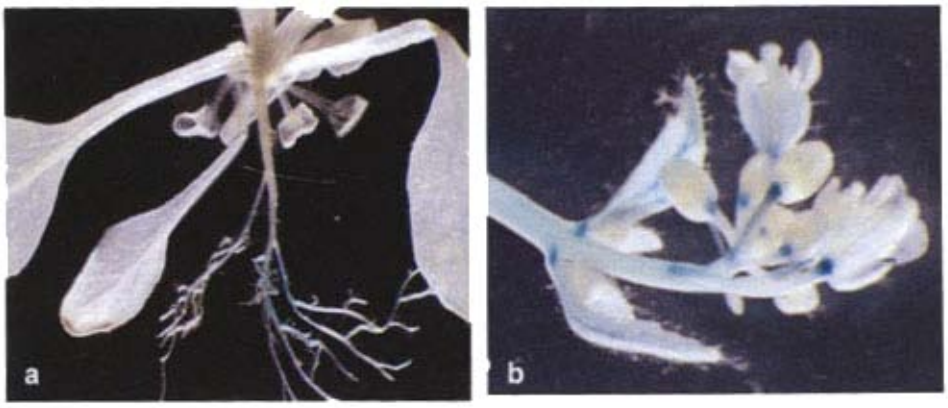
χ^2 at 0.05 = 3.84

Line No.	Basta R	Basta S	Total	$\chi^2$3: 1	P Value at 3:1	χ^2 15:1	P value at 15:1
1	**627**	**220**	**847**	**0.43**	**0.5120**	-	-
29	670	87	757	73.6	<0.0001	35.51	<0.0001
55	683	273	956	6.45	0.0111	811.84	<0.0001
57	**658**	**234**	**892**	**0.72**	**0.3961**	-	-
65	751	79	830	106.1	<0.0001	15.13	<0.0001
71	658	118	776	39.7	<0.0001	106.23	<0.0001
89	**663**	**195**	**858**	**2.36**	**0.1245**	-	-
99	638	166	804	8.13	0.0044	284.40	<0.0001
129	708	135	843	36.3	<0.0001	137.17	<0.0001

Earlier, Lindsey *et al.* (1993) found that 50-60% of selected *Arabidopsis* primary transformants exhibited a 3:1 ratio for segregation kanamycin resistance, indicative of the presence of T-DNA at a single locus. The number of independent T-DNA insertions was calculated to be ~1.3 per transformant in these lines. Similar results have been obtained in *Arabidopsis* T-DNA tagged lines, generated by earlier workers (Feldmann and Marks 1987; Bechtold *et al.*, 1993).

4.5 SCREENING FOR GUS EXPRESSION

More than 250 Basta resistant T_2 lines were screened for tissue specific expression of GUS in plants of individual lines at different stages of development. It is to be pointed out that expression of GUS only in different plant parts at some stages of plant development was looked for and an extensive screening of GUS under a variety of environmental conditions or responding to the presence of any chemical inducer was not carried out. A more extensive survey of these lines might lead to the identification of some new genes and promoters. Several mutants with reporter gene expression in different tissues were identified from the pCS1 promoter trap lines. Approximately, 21% of the plants showed GUS expression in different tissues. The frequency obtained by us is much higher than the low frequencies of GUS expression obtained by Richardson and co workers (6%) and Mollier and co workers (5-8 %) (Richardson *et al.*, 1998 and Mollier *et al.*, 1995). Both the groups used the pGKB5 promoter trap vector to generate their mutagenised populations. However, the results obtained by us are comparable to the results obtained by Koncz *et al.* (1989) using promoter trap vector, pPCV621where 30 % of the T-DNA transformants in *Arabidopsis* expressed a promoter-less *aph*(3') II gene conferring resistance to kanamycin. Several lines showed interesting GUS expression patterns e.g. a line expressing intense GUS activity in all tissues in the seedling stage was identified (Fig. 9c). Fig. 9b shows a mutant with GUS expression specifically at the bases of siliques and flowers. Several other lines showed GUS expression patterns in patches in siliques and buds



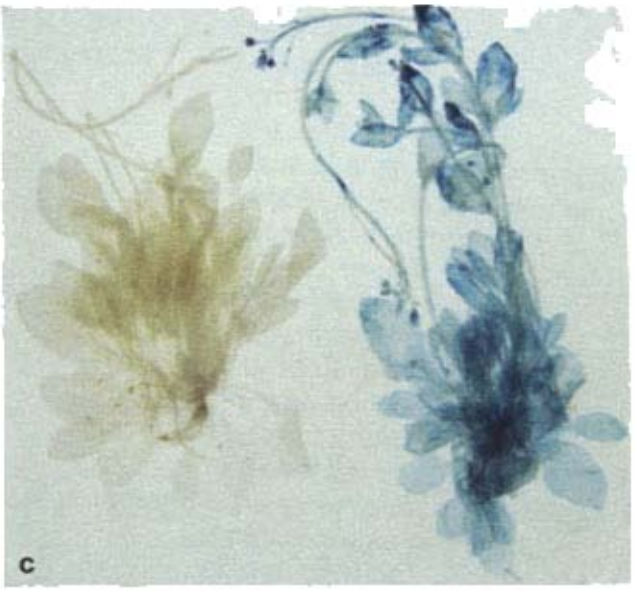


Fig. 9 Mutants showing GUS expression

a. M-57 root-specific

b. M-65 Bases of flowers in stalks and in veins

c. M-29 Constitutive in comparison to wild type

(Table 6). One line (M-57) that segregated in 3:1 ratio for Basta resistance in T_2 showed GUS expression specifically in roots (Fig. 9a).

Table 6. GUS Expression pattern in T_2 lines

SL. NO.	Mutant No.	Expression pattern of GUS
1	7	Stem portion
2	21	Base of the flowers and siliques
3	23	Shoot and stalk
4	29	Constitutive expression
5	50	Flowers siliques, stalk and shoot
6	52	Flowers and siliques
7	65	Shoot specific (Vascular Tissues in all veins)
8	69	Base of the siliques and stalk
9	81	Shoot specific
10	89	Siliques and shoots
11	99	Shoot
12	103	Siliques base (stalk)
13	111	Constitutive expression
14	120	Siliques base
15	125	Flowers, stalk of siliques and leaves
16	129	Siliques flowers and leaves
17	136	Constitutive expression
18	140	Siliques specific (at the base)
19	143	Constitutive expression
31	224	Pedicel and flowers
32	236 (A)	Constitutive expression
33	246	Flower, buds stem and stalks of siliques
34	254	Base of the siliques and stalks
35	270	All over the siliques wall (Weak expression at the gynoecium's wall, sepals and petals) mid rib and veins of leaves
36	**57**	**Root-specific**

Approximately, 30% of the T_2 mutant lines segregated for a 3:1 ratio for Basta resistance in the population. However many such lines did not show any GUS expression and in some of the cases where they did show GUS expression it was difficult to get to the gene because of the complicated nature of T-DNA integration in the form of tandem or inverted repeats or even the presence of vector backbone. A collection of 1,50,000 T_1 seeds is yet to be screened. The remaining seed collection has a potential to produce ~3000 independent primary transformants as

calculated from 2% average frequency of transformation. This population could prove to be good source for functional genomic studies. It can also be used for reverse genetic screens using left and/or right border/plant DNA junction specific PCR primers or the search for knock outs in specific genes using PCR based strategies. The population generated, can also be used in extensively screening various tissues at different developmental stages and also by monitoring reporter gene expression in response to various environmental cues. It can thus serve as a potential source for a variety of mutants.

4.6 A MUTANT SHOWING GUS EXPRESSION IN THE ROOTS

The mutant M-57 that showed root-specific GUS expression was selected for detailed analyses. The T_2 population fitted with the expected Mendelian segregation ratio of 3:1 for Basta resistance (Table 5), suggesting that the mutant has a single insertion. The GUS expression pattern of this line was studied thoroughly. GUS expression was found mainly in vascular regions of roots. The expression level was however very high in root tips. No GUS expression was observed in root hairs (Figs 10 a-d). GUS expression was not observed in any other plant part, however very rarely (in about 1 in 30 plants observed) a faint GUS expression was observed in some shoot tips of a plant. Histochemical assay for GUS revealed intense blue color in roots of young seedlings and faint blue in roots of older seedlings, suggesting that the expression of GUS in roots was very high in young seedlings but decreased as the seedling grew. In order to further confirm the presence of a single insert 500 T_2 plants were germinated and 375 Basta resistant ones selected. All the Basta resistant

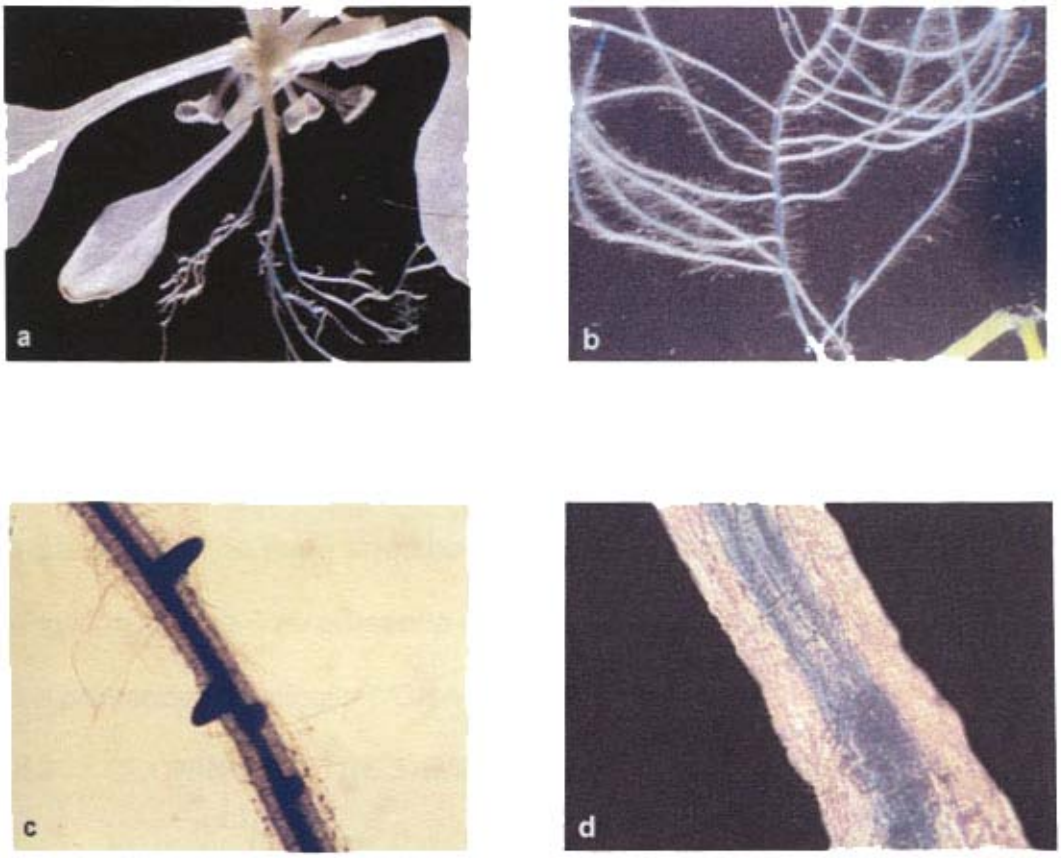


Fig. 10 Detailed GUS expression analysis of M-57

a. Mutant seedling showing root-specific expression

b. A closer view of the roots showing intense blue color at the root tips and all along the vascular tissue

b. 20X Magnification GUS expression restricted to vascular tissue and very high expression in emerging lateral root tips

d. 40 X Magnification clearly shows that GUS expression is limited to vascular tissue

plants showed GUS expression in roots thus proving co-segregation of *gus* A gene with the *bar* gene.

4.7 SOUTHERN HYBRIDIZATION

In order to confirm the inheritance and copy number of T-DNA in M-57, a Southern blot analysis with M-57 DNA digested with *Hin*d III and *Nde* I was carried out. The probe used was the 1.2 kb *gus* amplicon amplified from the pCS1 plasmid DNA. A 3.8 kb band with *Hin*d III and a 10 kb band with *Nde* I was obtained. As *Nde* I is not present in the T-DNA region it was expected that the band size should be greater than 7kb (size of the T-DNA) (Fig. 11a and b). The presence of a single band in both the lanes confirmed the presence of a single T-DNA insert in the mutant.

4.8 CLONING OF THE T-DNA FLANKING SEQUENCE

In order to clone the T-DNA flanking sequence of this mutant, inverse PCR was carried out. DNA was restricted with *Nde*I because the restriction site for this enzyme was not present in the T-DNA. The primers *TR3* (near Right T-DNA border) and *LB2* (near Left T-DNA border) as depicted in Fig. 12 were used for inverse PCR. A 1.5 kb amplicon was obtained (Fig. 13) which was cloned in pGEM T Easy (TA Cloning Vector). The positive clones were confirmed by *Eco*R I restriction (Fig. 14), which yielded a 1.0 kb and a 0.5 kb restriction product signifying the presence of a *Eco*R I site in the insert. The fragment was then sequenced from both the ends (A.9). Plasmid rescue to clone the flanking sequence did not succeed although it was possible to rescue a smaller plasmid from the complete pCS1. Inverse PCR was performed on DNA sample that was digested with *Nde*I and self-ligated, which showed that a circular plasmid was indeed

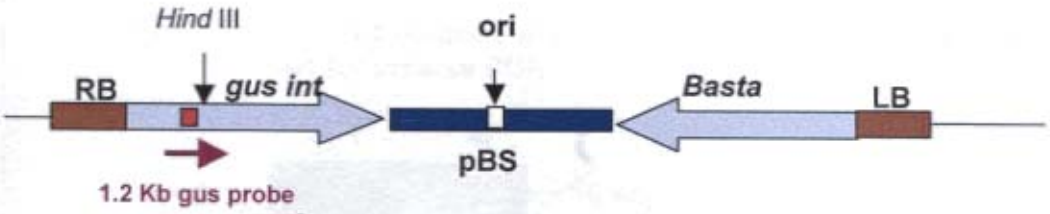


Fig. 11a Schematic diagram of the T-DNA showing position of probe

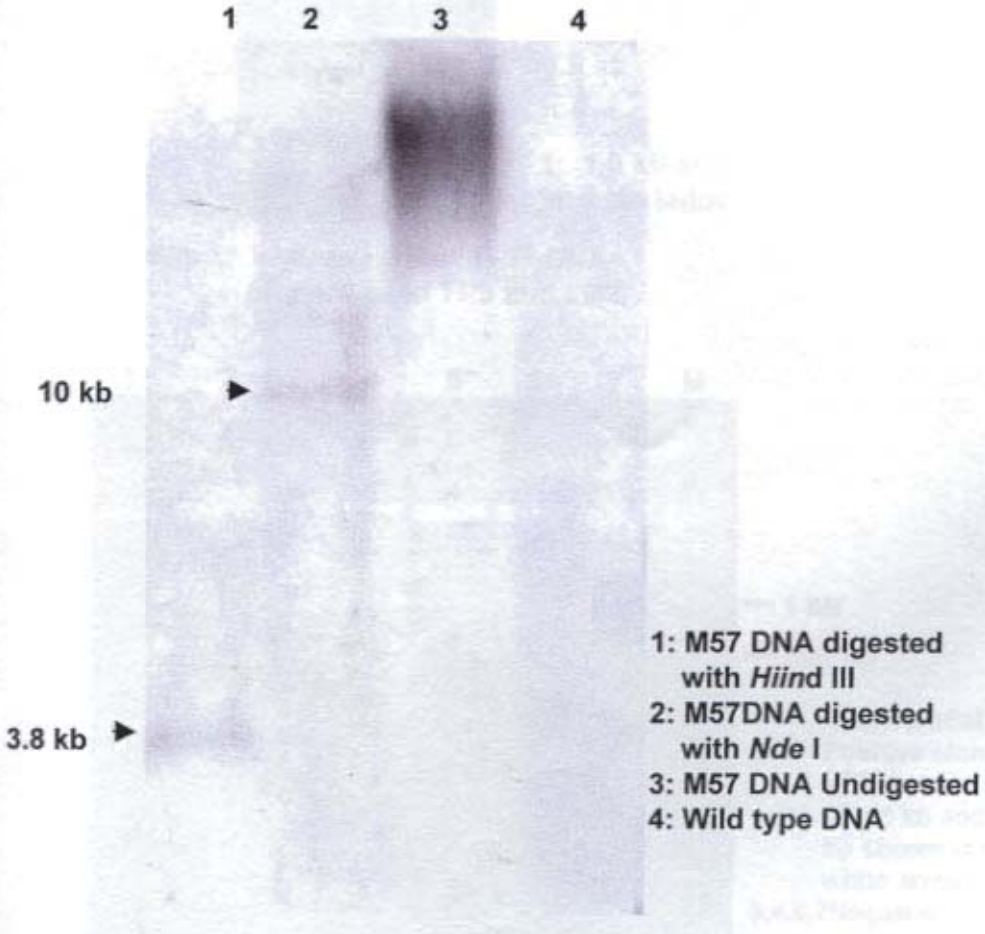


Fig. 11b Southern hybridization of M-57 DNA with *gus* probe

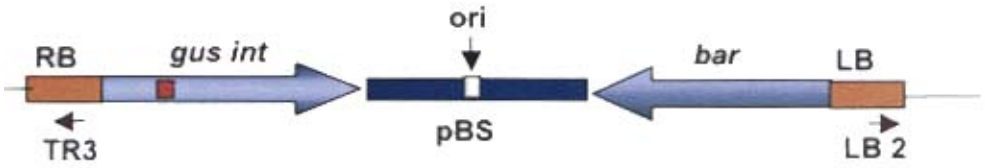


Fig. 12 Schematic diagram of the T-DNA showing the two primers used for inverse PCR

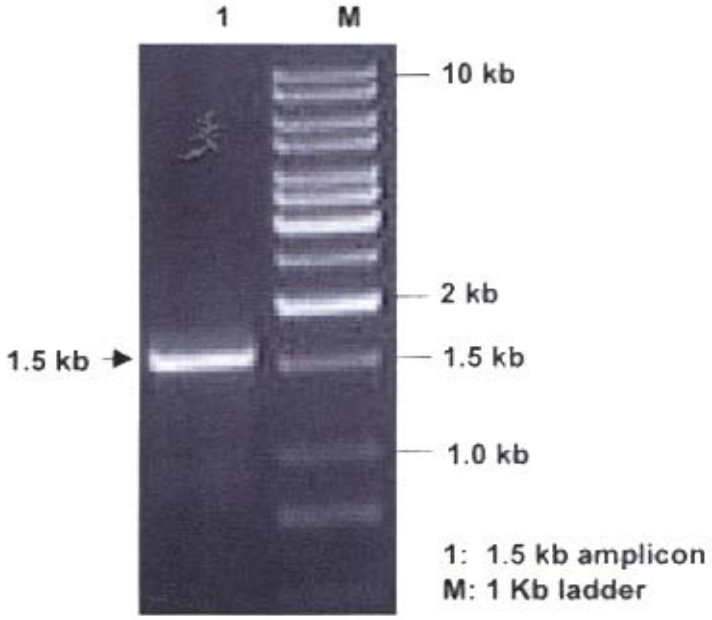


Fig. 13 Inverse PCR of M-57 DNA primers used TR3 and LB2

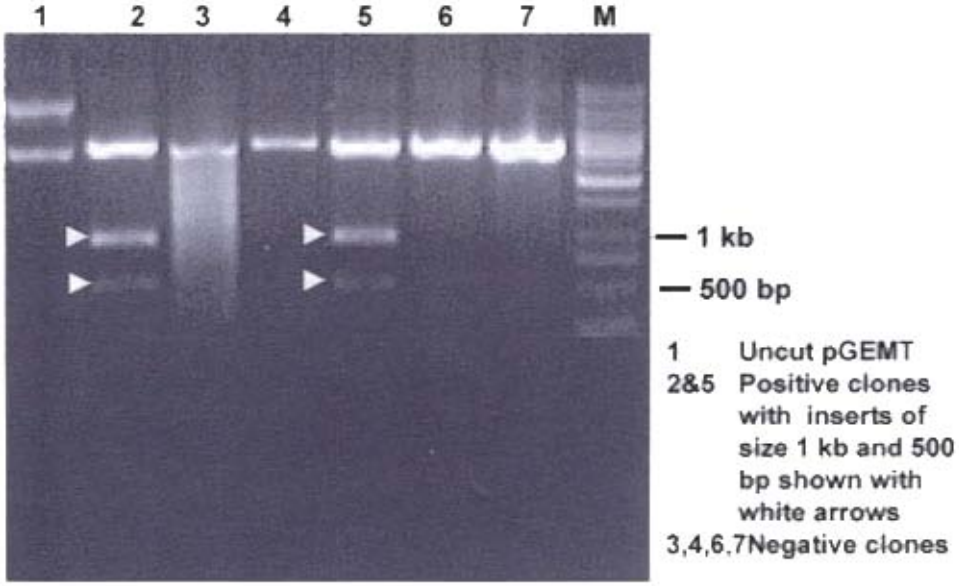


Fig. 14 Confirmation of pGEM-T derived positive clones containing the 1.5 kb insert

formed. It is quite possible that the pBluescript origin of replication present in the T-DNA region became non-functional following integration in the plant genome. Weigel *et al.* (2000) have reported that the isolation of T-DNA flanking sequences in T-DNA knockout mutants by plasmid rescue was difficult due to the complicated nature of T-DNA insertions.

4.9 COMPARISON OF THE SEQUENCE WITH THE *ARABIDOPSIS* GENOME

The sequence obtained was aligned with the *Arabidopsis* genome at the www.arabidopsis.org website (TAIR Website) and compared with the whole genome including BAC clones. The T-DNA insertion was in the intergenic region between two loci At4G13940 and At4G13930 i.e. in the promoter area between two genes. The locus At4G13940 has been already annotated as a S-Adenosyl Homocysteine Hydrolase (SAHH) while the locus At4G13930 as Serine Hydroxy Methyl Transferase (SHMT). These loci transcribe in opposite directions and the insertion is in the upstream region of both the genes ~ 1kb from the translation start site of SAHH and 3.6 kb from the translation start site of SHMT (Figs 15a and b). The latter is expressed maximally in roots in a circadian rhythm (Mc Clung *et al.* 2000), while the former has been reported to be an embryo defective gene (EMB 1395) (Tzafrir *et al.*, 2003 and www.seedgenes.org).

4.10 *IN SILICO* RESTRICTION ANALYSIS OF INSERTION SITE

An *in silico* restriction analysis of sequences around the insertion site confirmed the signals obtained by Southern hybridization earlier (Figs 16 and 11b). DNA digested with *Hin*d III yielded, as expected, a 3.8 kb band when probed with the 1.2 kb *gus* probe and that digested with *Nde* I

taatatatttttatatcatatttattatttcgatcgatttccaatagaaacaatttcgataaaattacatt
tttttgtaaataatatataaatttttttgtagattacattaatagaactgctgaaaatggtctttaaatcctct
tcctcttcaacatttttccaaatttgacattttgactccgaatctcacttataagttgtaataataatgttca
ttagtgcattagtttcctcttcgttgcccaacaaaagaataagctatcattgacagttttatgatattcgga
gttccaaaatcatttttcgaaactcggcctttaaggcatttggatcaaaacttcggtgcatatactaaaaac
atatggatatatcgggttgagtcgcctcgtagaattcctcgagaagatatatattttgacttggtgctggcca
tttgtaagaatagcataagccaccagtcaccaagattattcatcgtaaattgtaacaaactcaccattatttg
gagagcccaaagatactactcaatcttcttaccataaacttttgtcactgtcttataagctataagattactg
gtctacttaactcagaaatatttatgtcctttgtagccaaaagctacactgaaattgacacaactcgagtgtt
gtccttttctggtcgattgaatagaatcgaatgtcttaatccagtaccctccagcttttatttcgtgtaatttt
attttccaaacctaccactaccagtttcataactctcgaataaatttatcaaatagtcttttgagtgctcaaa
gtcttgggataataaatggtcagtgctatgtatcacccggatgtgaaacattatgggtggagatagactatta
taaatttattgaaatatacgattgttactcgtttaatagcaaaagtagtacaatgtatatagtttctatcgag
aacaagatctatttaaaattcgaaaagtacatttaaaattcataaacatataaagatagtaacatgttagatc
gcatagtaccacccaaaacaagaaaaaagaaacgcacatcgccacataattgctatgattctcactgtcggctg
ctttgaaatattacgattcttttggtaaatcacacaacataatatataattacaataaatatatatatactaaag
tataattaatatatattaatacccacattgtttaattctgttttgatcttttaagatcagtcagatccaccgacg
ttcctacacgcgcaggtccagatccaaacag..T-DNA RB (Adjoining sequence)

cacacacacacacacaatgccactagtgtaaatgcttggtggctattgcatttgcacc(Deleted portion
in the T-DNA mutant)

T-DNA Left Border(Adjoining sequence) tattgatactctttcttcaaaaacaagttattgtttt
tattttcaacccaactttaatacggattcatactgggatttaggtgttaaatctgataatttaggtttgaataag
ttgtatatttgtttctttgattaaaaaaagaacctatatatatacaaaaataaataaaaagttctagatttcaat
tttcgtatatagcgggttgaattgtctatttaatatgaaaattgacggatcttataaacaaaatgttctgaaa
tatgtaaaaggatttagccaaagttaaccaaaaaaaaaaaacaaacagaaaagtcacattcacatgtcgtggtag
atctaaggcattaatttagaaatatgtcgttacaataagcggagaacatgggacgtttctcgtggtccaatcaga
cgaacgagatctcataaattaaatgacttcagacgagggaattcatggcagaatgataatgcaacttaagtgact
ttagagtgaaaatgatacgagaacaatgcataatccatatgaccgttgagtgagtgataccattagcgcgatacа
agcgggactataaactgatctagattgtttttcttgggaaaaaatgttacaaattttaaatatgtagtttgaatt
gttaaaccaagattcaacagaaatatacccgtaaataaacaacagttgataatagtcatcgaaaagatatcaactg
attcttcacttgggctactgtgacggcccgttagggttctcaatataagtcaatаactacgatctacgattcact
gaaacaaataaaacacagccacgtgtccacccctcccacatcaccgtccgatctaccccacgacaagcttacaaca
cgggtcataccgctcgtgcagcgtgttccgtcatccacgggattacaacttctaccagatccaccaaaccctcaa
aacaatctgaaccgttcatttcattttgacctcatctatatattctctgtcactccctttctcttctc

SAHH starts

Fig. 15a The sequence flanking the T-DNA insertion LB Flanking sequence is shown in green RB Flanking sequence is shown in pink

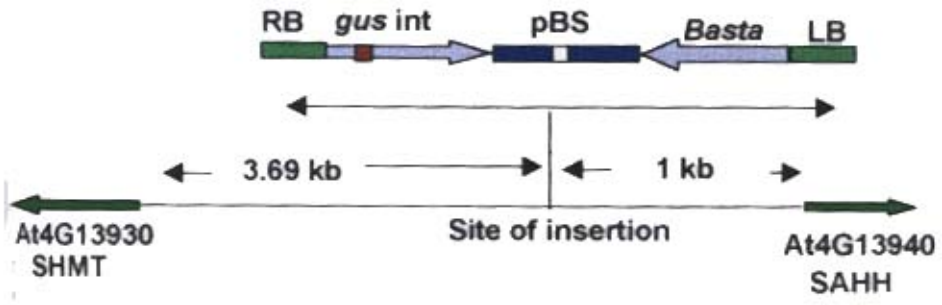


Fig. 15b Schematic diagram showing the position of the T- DNA insertion in the intergenic region between two genes

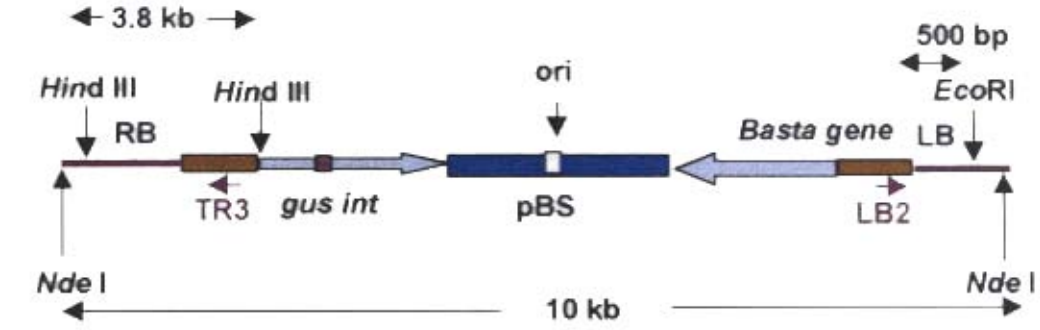


Fig. 16 Schematic diagram of T-DNA in the *Arabidopsis* genome depicting the presence of restriction sites and the two primers used for inverse PCR

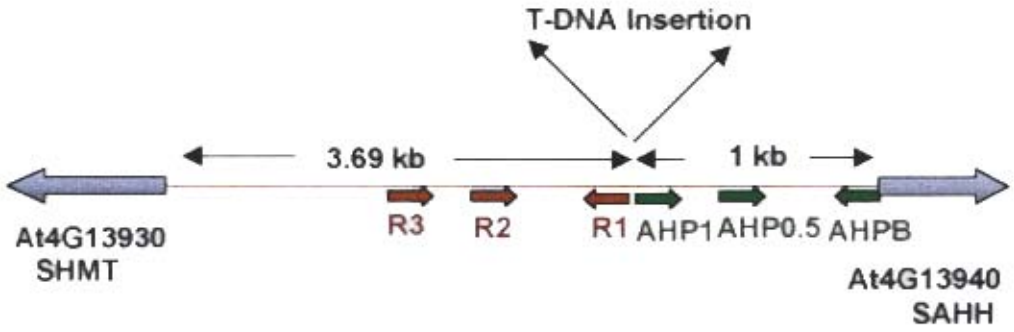


Fig. 17 Schematic diagram depicting the primers designed to clone upstream regulatory regions. The primers upstream of the T-DNA insertion are shown in red while those upstream of SAHH gene are shown in green

yielded a 10 kb band. This is because of the presence of a *Nde* I site about 1 kb downstream of the T-DNA left border and another one about 2 kb upstream of the T-DNA right border. Similarly one *Hin*d III site was present 1.6 kb upstream of the T-DNA right border and another one about 2.2 kb downstream of the T-DNA right border in the pCS1 plasmid. The presence of an *Eco*R I site about 500 bp downstream of the T-DNA left border explained the occurrence of two restriction products of size 1 kb and 500 bp found in positive clones as described earlier (Figs 16 and 14).

4.11 CLONING OF UPSTREAM REGULATORY REGIONS IN A BINARY VECTOR

In order to clone the 5' regulatory regions, upstream to *gus* gene, two different sets of three primers were designed as depicted schematically in Fig. 17 (details regarding primers are mentioned in Materials and Methods, section 3.14 and A.8). The primers R1, R2 and R3 were used to amplify two different fragments of 905 bp (R1 and R3) and 452 bp (R1 and R2) upstream of the T-DNA insertion. The primers AHB, AH 0.5 and AH1 were used to amplify two different fragments 1180 bp (AHB and AH1) and 540 bp (AHB and AH0.5) upstream of gene SAHH (Fig. 18). The amplicons were cloned in pDRIVE (UA cloning vector) and the authenticity of the insert was confirmed by sequencing the clone. After restriction of pDRIVE derived positive clones with *Bam*H I and *Sal* I the desired inserts were released, which were subsequently cloned into pBI101 upstream of the *gus* gene. All the resulting pBI101 derived positive clones were confirmed by digestion of the plasmid DNA with *Bam*H I and *Sal* I, that released the inserts of sizes 542 bp, 452 bp, 1180 bp and 905 bp as shown in Fig.19. The constructs carrying the two different fragments (452 bp and 905 bp)

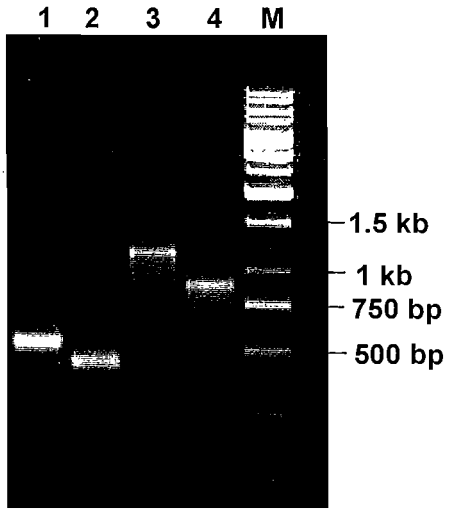


Fig. 18 PCR amplification from wild type *Arabidopsis* genomic DNA showing
Lane1: 542 bp with AHPB and AHPB 0.5
Lane2:452 bp with R1 and R2
Lane3: 1180 bp with AHPB and AHPB1
Lane4: 905 bp with R1 and R3
Lane M: 1 kb ladder

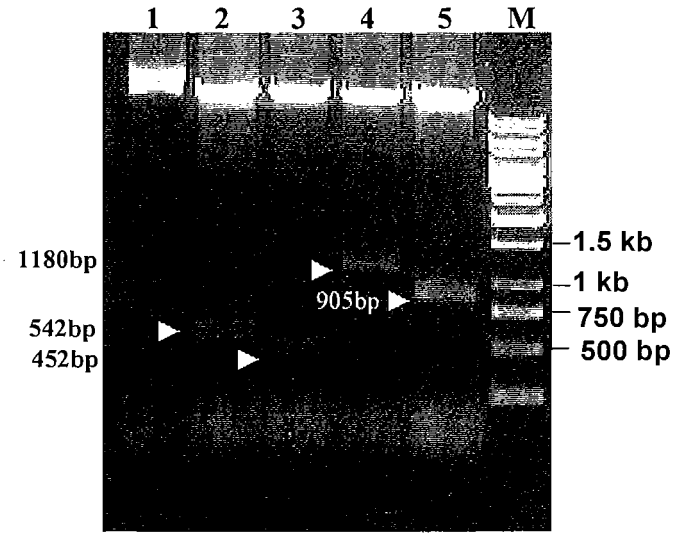


Fig. 19 The four positive pBI clones digested with *Bam*H I and *Sal* I
Lane 1 Uncut pBI 101
Lane 2 pCS2-*Bam*H I and *Sal* I
Lane 3 pCS3-*Bam*H I and *Sal* I
Lane 4 pCS5-*Bam*H I and *Sal* I
Lane 5 pCS4-*Bam*H I and *Sal* I
Lane M 1 kb ladder

upstream of the T-DNA insertion were designated pCS2 and pCS3 respectively (Fig. 20a) and those carrying the fragments upstream of SAHH gene (542 bp and 1180 bp) were designated pCS4 and pCS5 (Fig. 20b). Table 7 summarizes the different pBI derived constructs.

Table 7. The pBI-derived constructs

S No.	Primer pair	Construct	Fragment cloned
1	R1 and R2	pCS2	452 bp
2	R1 and R3	pCS3	952 bp
3	AHB and AH 0.5	pCS4	542 bp
4	AHB and AH1	pCS5	1180 bp

The constructs were mobilized into *Agrobacterium* strain GV3101 and positive colonies identified by colony PCR. The regions upstream of the gene SAHH were also cloned for three reasons: (i) the expression pattern of this gene was not known, although its mutant phenotype was known, (ii) the main reason was to use this sequence as a standard promoter as it lies immediately upstream to a gene so that it could be used to compare the expression pattern of the sequence upstream to the T-DNA insertion, which lies in an intergenic portion and (iii) since this gene lies in the same orientation as the *gus* gene in the mutant, it would be interesting to check for its GUS expression pattern too.

4.12 *ARABIDOPSIS* TRANSFORMATION WITH pBI101 DERIVED CONSTRUCTS AND HISTOCHEMICAL ANALYSIS OF GUS

Arabidopsis plants were transformed with *Agrobacterium* strain GV3101 containing the pBI-derived constructs by floral dip method. The seeds were selected on MS medium containing kanamycin. A total of 10 T_1 transformants were obtained with construct pCS2 and 15 with pCS3. With

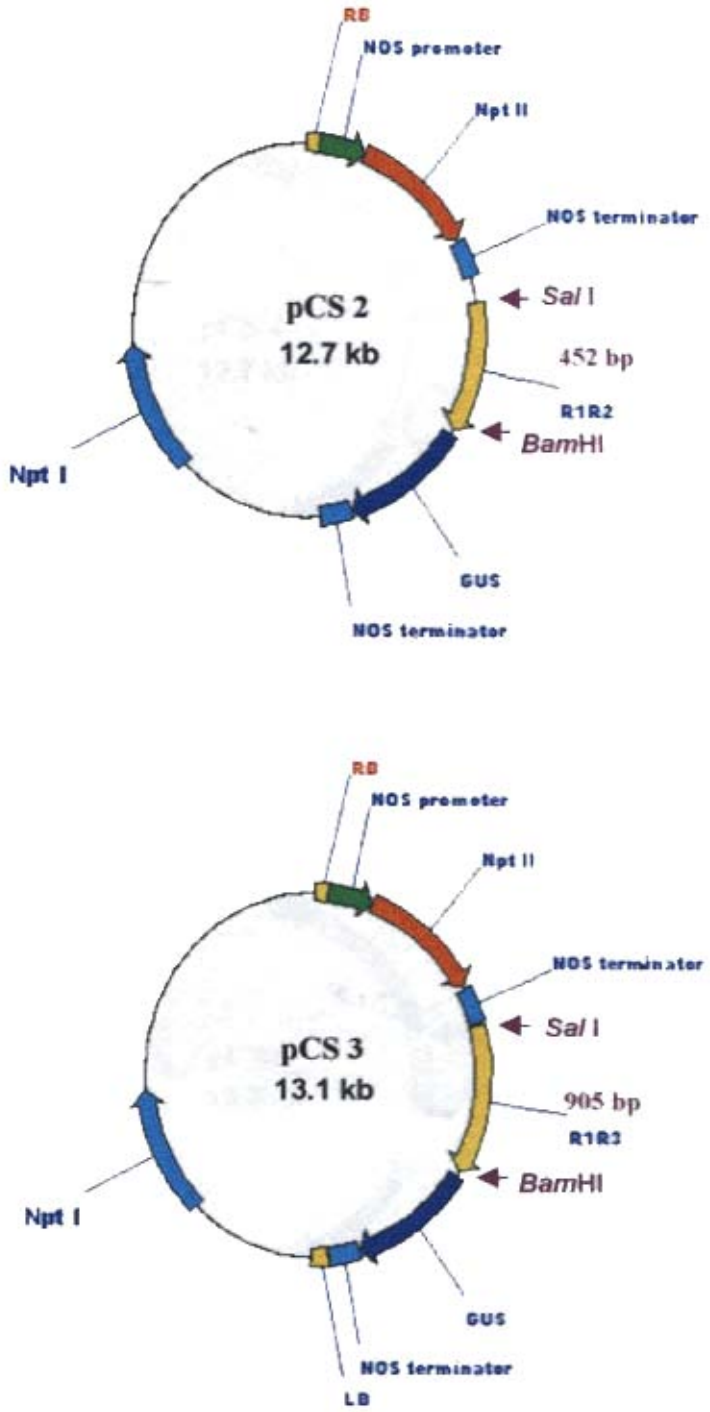


Fig. 20a pBI derived constructs carrying the 452 bp (pCS2) and the 905 bp (pCS3) fragments upstream to the T-DNA insertion at the *Bam*HI and *Sal* sites of pBI101

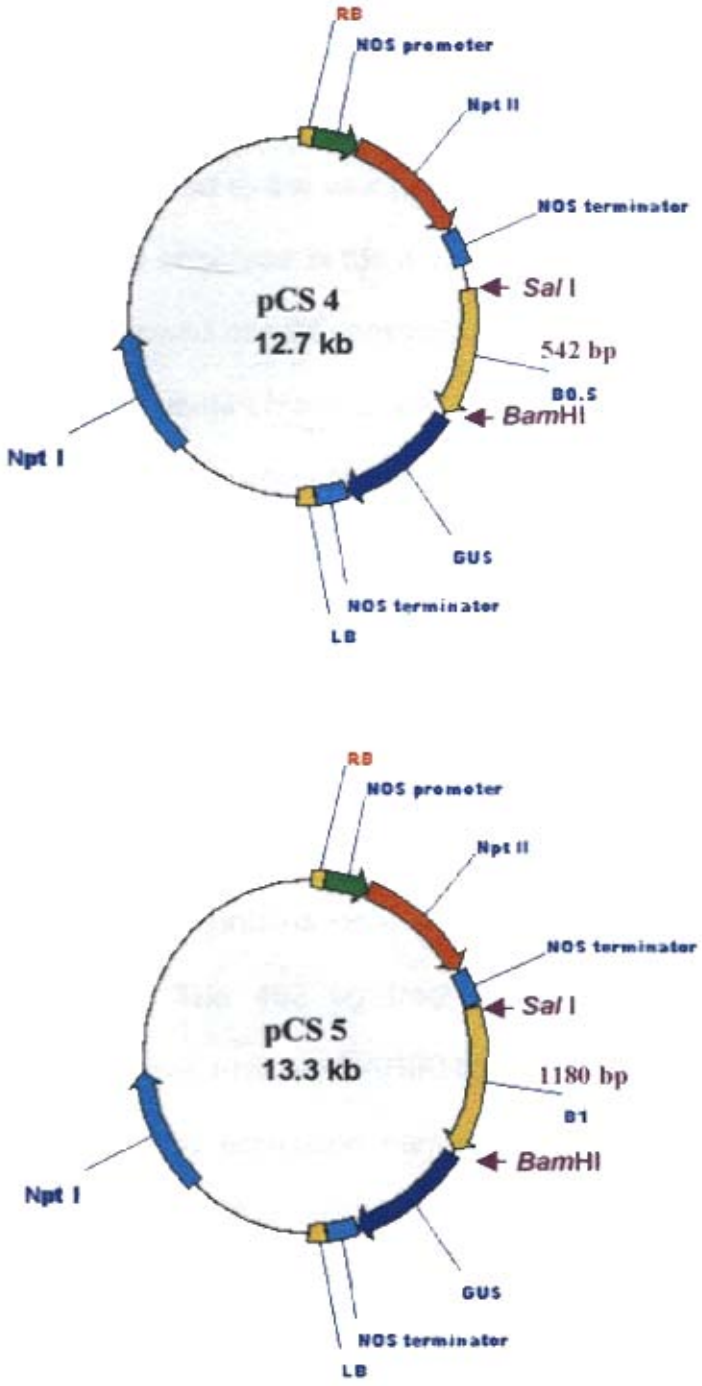


Fig. 20b The new pBI derived constructs carrying the 542 bp (pCS4) and the 1180 bp (pCS5) fragments upstream to the SAHH gene at the *Bam*HI and *Sal*I sites of pBI101

pCS4 and pCS5 a total of 10 and 15 T_1 transformants respectively were selected. The GUS expression patterns of T_2 plants of the selected T_1 transformants obtained with pCS2 and pCS3 showed root-specific expression, that too, limited to the vascular tissues and root tips similar to the expression of GUS observed in the mutant M-57 (Fig. 21), while those of pCS4 and pCS5 showed almost constitutive expression in the seedlings (Fig. 22). The transformants obtained with pCS2 and pCS3 did not show GUS expression in any other plant part while the transformants of the other two constructs showed varied expression patterns in different plant parts at different stages of development especially in leaves.

The 452 bp portion (amplified using primers R1 and R2) upstream to the T-DNA insertion, although occurring in an intergenic region is independently able to drive root specific expression, which confirms the fact that the 452 bp region contains *cis*-acting regulatory elements for root-specific expression. The 452 bp fragment and the 540 bp portion (amplified using primers AHB and AHB0.5) have been submitted to the Genbank database vide accession numbers AY601849 and AY601850 respectively (A.10 and A.11).

4.13 *IN SILICO* CHARACTERIZATION OF THE UPSTREAM REGULATORY SEQUENCES

The sequences upstream of the SAHH and those upstream of exact insertion site were analyzed for the presence of *cis*-acting regulatory elements using a web based program Plant cis-acting regulatory elements (Plant CARE). Fig. 23 shows a schematic diagram of the promoter of the SAHH showing different motifs. An already reported root specific motif, as1 (Yamaguchi-Shinozaki and Shinozaki, 1994) is present 34 bp upstream of

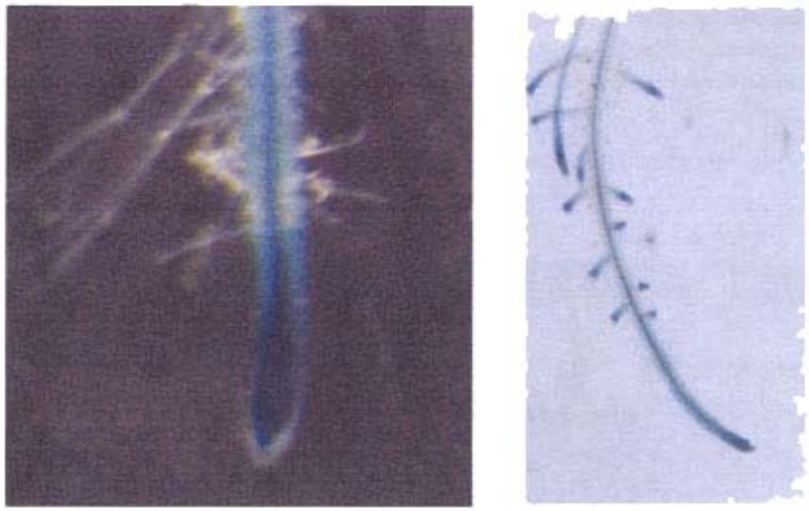


Fig. 21 GUS expression assay on seedlings transformed with pCS2 and pCS3. Note high GUS expression in root tips and GUS expression limited to vascular tissues in roots.

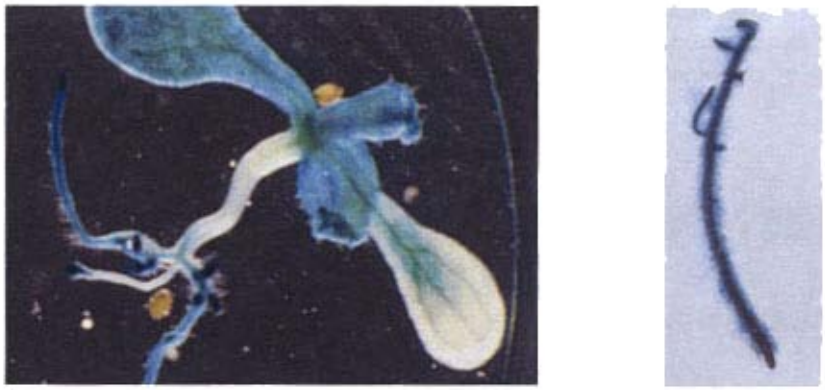


Fig. 22 GUS expression assay on seedlings transformed with pCS4 and pCS5. Note almost constitutive GUS expression and GUS expression all over roots

cgagtgataccattagcgcgatacaagcgggactataaactgatctagattgttttt
cttgggaaaaaatgttacaaattttaaatatgtagtttgaattgttaaaccaagatt
caacagaaatataccgtaaataaacaacagttgataatagtcatcgaaaagatatca
actgattcttcacttgggctactgtgacggcccgttagggttctcaatataagtcaa
taactacgatctacgattcactgaaacaaataaaacacagccacgtgtccaccctcc
cacatcaccgtccgatctaacccacgacaagcttacaacacgggtcataccgctcgt
gcagcgtgttccgtcatccacgggattacaacttctaccagatccaccaaaccctca
aaacaatctgaaccgttcatttcattttgacctcatctatatttctctgtcactcc
ctttctcttctc..... SAHH Gene

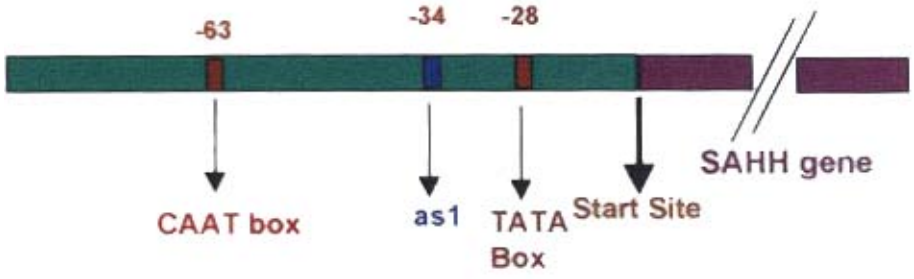


Fig. 23 Schematic diagram depicting the structure of the 504 bp promoter of the SAHH gene showing different *cis* acting regulatory elements. The upper panel shows the sequence

gctatgtatcacccggatgtgaaacattatgggtggagatagactattataaatttat
tgaaatatacgattgttactcgtttaatagcaaaagtagtacaatgtatatagtttcta
tcgagaacaagatctatttaaaattcgaaaagtacatttaaaattcataaacatataaa
gatagtaacatgttagatctgcatagtaccaccaaaacaagaaaaaagaaacgcacatc
gccacataattgctatgattctcactgtcggctgctttgaaatattacgattcttttgg
taaatcacacaacataatataattacaataaatatatatatactaaagtataattaata
taattaataccacattgtttaattctgttttgatcttttaagatcagtcagatccaccg
Acgttcctacacgcgcaggtccagatccaaacag .. Intergenic region

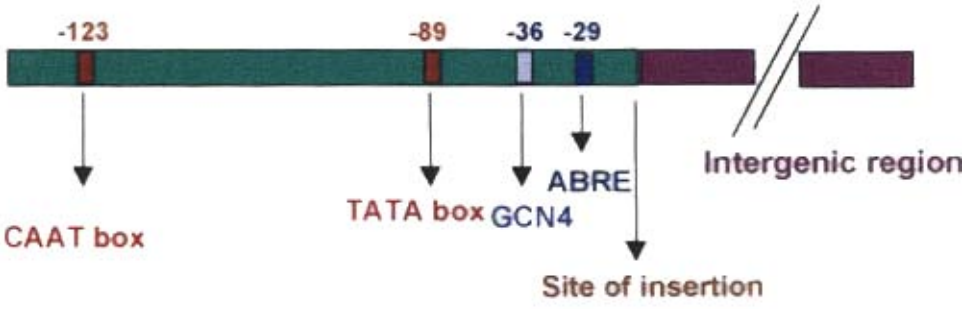


Fig. 24 Schematic diagram depicting the structure of the 450 bp fragment upstream of the T-DNA insertion showing different *cis* acting regulatory elements. The upper panel shows the sequence

transcription start site between the TATA and CAAT boxes. Although the TATA box of the SAHH promoter occurs at the usual position it doesn't conform to the consensus sequence, TATAAA. In comparison, the sequences upstream to the T-DNA insertion (Fig.24) has a TATA box 89 bases upstream of the site of insertion and a CAAT box 123 bases upstream of the site of insertion. It is to be noted that the TATA box conforms to the consensus sequence TATAAA. An abscisic acid responsive element (ABRE) and a motif for endosperm specific expression (GCN4) occur even before the occurrence of the first TATA box and no known root specific promoter element was found in this sequence. It is almost tempting to suggest that the 452 bp fragment could be a promoter buried in the vast intergenic region between the two genes, SAHH and SHMT, i.e. it could be a cryptic root-specific promoter. It is evident that the promoter got activated when placed adjacent to the *gus* gene. Cryptic elements are present in all organisms as inactive elements at their native locations in the genome but are activated when positioned in front of a gene (Fobert *et al.*, 1994 and Fourel *et al.*, 1992). Thus mutant populations generated through promoter trap vectors are more likely to reveal the presence of crypic promoters as indeed the one cloned and identified in the present study. Several cryptic promoters have been isolated in plants (e.g. Fobert *et al.*, 1994 and Mollier *et al.*, 2000). The constitutive cryptic promoter tCUP from tobacco (Foster *et al.*, 1999) has been analyzed and its structure and enhancer elements shown to be similar to other identified promoters of plant genes (Wu *et al.*, 2001 and Wu *et al.*, 2003). In this study, a cryptic root-specific promoter has been identified. A thorough

deletion analysis of the cryptic root-specific regulatory element could reveal the presence of some key root specific enhancer elements and also the core promoter area including the position of TATA boxes and initiator elements.

SUMMARY AND CONCLUSION

Promoter trapping is a powerful technique for the isolation of tissue specific promoters. Using this approach a cryptic root-specific promoter from *Arabidopsis* has been cloned and characterized. The salient points of the study are summarized below

1. A promoter-trap vector, pCS1, carrying a promoter less *gus* reporter gene fused to the right border of T-DNA, a plasmid origin of replication and a *bar* gene as a plant selectable marker was constructed.

2. The plasmid pCS1 was used to transform *Arabidopsis thaliana* ecotype Columbia. The transformants were selected on MS plates containing Basta and a population of *Arabidopsis* mutants has been generated. A part of this population has been screened to obtain 500 independent mutant lines.

3. Approximately, 30% of the T_2 mutant lines segregated in a 3:1 ratio for Basta resistance in the population. 20% of the mutants, showed GUS expression in different tissues. Among such lines one, M-57, showed GUS expression specifically in roots. Selfed T_2 plants of M-57 segregated in a 3:1 ratio for Basta resistance. The GUS expression was also found to co-segregate with Basta resistance. In M-57 GUS expression in roots was confined mainly to vascular regions of roots and was very pronounced in root tips.

4. Southern hybridization analysis of M-57 genomic DNA with *gus*-specific probe confirmed the presence of a single T-DNA insert.

5. The Genomic DNA flanking the T-DNA was cloned by inverse PCR and sequenced.

6. The sequence was aligned with the *Arabidopsis* genome and it was found that the T-DNA had inserted in an intergenic portion in the chromosome no. 4 between two genes, At4G13930.1 annotated as Serine hydroxy Methyl transferase (SHMT) and At4G13940.1 annotated as S-Adenosyl Homocysteine Hydrolase (SAHH). These two genes are transcribed in opposite directions and the insertion was found to be 1kb upstream of the ATG of SAHH and 3.7kb from the ATG of SHMT. The direction of transcription of *gus* was in the same direction as SAHH gene.

7. Two DNA fragments upstream of the T-DNA insertion, 905 bp and 452 bp and two others upstream of gene SAHH, 1180 bp and 540 bp were cloned upstream to the *gus* gene in binary vector, pBI101 and four corresponding pBI derived constructs, pCS3 (905 bp); pCS2 (452 bp); pCS5 (1180 bp) and pCS4 (540 bp) were generated respectively. These constructs were used to transform wild type *Arabidopsis* plants and the transformants selected on MS plates containing kanamycin. The T_2 seedlings of these resistant lines were analyzed for GUS expression.

8. The pattern of expression of GUS in transformants obtained from pCS2 and pCS3 were faithfully reproduced just like in the mutant M-

57. The transformants from the other two constructs, pCS4 and pCS5, carrying the sequences upstream of SAHH showed almost constitutive GUS expression.

9. The 450 bp fragment upstream of T-DNA insertion is independently capable of driving root-specific GUS expression, since this fragment is in the intergenic region it is likely that it is a cryptic root specific promoter.

10. Both the sequences, the fragment upstream to the T-DNA insertion and that upstream to SAHH gene have been assigned a promoter function and submitted to the Genbank database (vide accession numbers AY601849 and AY601850.

Promoter trapping is a very powerful technique to isolate tissue-specific promoters. In the present study a cryptic root-specific promoter was isolated from an *Arabidopsis* promoter trap population and the regulatory sequence shown to drive reporter expression specifically in roots has been identified. A further detailed analysis would reveal all the features of this novel cryptic root-specific promoter.

REFERENCES

Albani, D., Sardana, R., Robert, L.S., Altosaar, I., Arnison, P.G. and Fabijanski , S.F. 1992. A *Brassica napus* gene family which shows sequence similarity to ascorbate oxidase is expressed in developing pollen. Molecular characterization and analysis of promoter activity in transgenic tobacco plants. *Plant J.* **2**: 331-42.

Alonso, J.M., Stepanova, A.N., Leisse, T.J., Kim, C.J., Chen, H., Shinn, P., Stevenson, D.K., Zimmerman, J., Barajas, P., Cheuk, R., Gadrinab, C., Heller, C., Jeske, A., Koesema, E., Meyers, C.C., Parker, H., Prednis, L., Ansari, Y., Choy, N., Deen, H., Geralt, M., Hazari, N., Hom, E., Karnes, M., Mulholland, C., Ndubaku, R., Schmidt, I., Guzman, P., Aguilar-Henonin, L., Schmid, M., Weigel, D., Carter, D.E., Marchand, T., Risseeuw, E., Brogden, D., Zeko, A., Crosby, W.L., Berry, C.C. and Ecker, J.R. 2003. Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science*. **301**: 653-657.

Alvarado, M.C., Zsigmond, L.M., Kovacs, I., Cseplo, A., Koncz, C. and Szabados, L.M. 2004. Gene trapping with firefly luciferase in *Arabidopsis*. Tagging of stress-responsive genes. *Plant Physiol.* **134**: 18-27.

Anandhan, S. 2004. Isolation and charactortization of seed specific promoter in cowpea (*Vigna unguiculata* (L.) Walp.) *PhD Thesis. IARI, New Delhi.*

Anjanasree, K.N. 2003. Cloning and Characterization of Fruit-Specific Promoters from Tomato (*Lycopersicon esculentum* Mill.) *PhD Thesis. IARI, New Delhi.*

Arabidopsis Genome Initiative. 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana. Nature.* **408**: 796-815.

Ascencio, J. 1997. Root secreted acid phosphatase kinetics as a physiological marker for phosphorous deficiency. *J. Plant Nutr.* **20**: 9-26.

Atkinson, H.J., Urwin, P.E. and McPherson, M.J. 2003. Engineering plants for nematode resistance. *Annu Rev Phytopathol.* **41**: 615-639.

Azpiroz-Leehan, R. and Feldmann, K. A. 1997. T-DNA insertion mutagenesis in *Arabidopsis:* going back and forth. *Trends Genet.* **13**: 152-156.

Bade, J., van Grinsven, E., Custers, J., Hoekstra, S. and Ponstein, A. 2003. T-DNA tagging in *Brassica napus* as an efficient tool for the isolation of new promoters for selectable marker genes. *Plant Mol. Biol.* **52**: 53-68.

Balter, M. 1999. PLANT SCIENCE: Data in Key Papers Cannot Be Reproduced. *Science.* **283**: 1987-1989.

Banks, M.K., Kulakow, P., Schwab, A.P., Chen, Z. and Rathbone, K. 2003. Degradation of crude oil in the rhizosphere of Sorghum bicolor. *Int J Phytoremediation.* **5**: 225-234.

Barthels, N., Van-der-Lee, F. M, Klap, J., Goddijn, O. J. M., Karimi, M., Puzio, P., Grundler, F. M. W., Ohl, S. A., Lindsey, K., Robertson, L., Robertson, W. M., Montagu, M. V., Gheysen, G., Sijmons, P. C. 1997. Regulatory sequencs of *Arabidopsis* drive reporter gene expression in nematode feeding structures. *Plant Cell.* **9**: 2119-2134.

Baroux, C., Blanvillain, R., Moore, I.R. and Gallois, P. 2001. Transactivation of BARNASE under the AtLTP1 promoter affects the basal pole of the embryo and shoot development of the adult plant in *Arabidopsis*. *Plant J.* **28**: 503-15.

Bechtold, N., Ellis, J. and Pelletier, G. 1993. *In Planta Agrobacterium*-mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. *C.R. Acad. Sci. Paris, Sci. Life Sci.* **316**: 1194-1199.

Benfey, P.N., Linstead, P.J., Roberts, K., Schiefelbein, J.W., Hauser, M.T. and Aeschbacher, R.A. 1993. Root development in Arabidopsis: four mutants with dramatically altered root morphogenesis. Development. **119**: 57-70.

Bent, A.F. and Clough, S.J. 1998. *Agrobacterium* germline transformation: transformation of *Arabidopsis* without tissue culture. In Plant Molecular Biology Manual (Gelvin, S.B. ed.). Netherlands: Kluwer Academic Publishers, B7 1-14.

Berg, C.M. and Berg, D.E. 1987. Uses of transposable elements and maps of known insertion. In *Escherichia coli* and *Salmonella typhimurum*. Cellular and Molecular Biology'. (Ed. F.C. Neidhart.) 1071-1109.

Bird, C.R., Smith, C.J.S., Ray, J.A., Moureau, P., Beran, M.W., Bird, A.S., Hughes, S., Morris, P.C., Grierson, D. and Schuch, W. 1988. The tomato polygalacturonase gene and ripening-specific expression in transgenic plants. *Plant Mol. Biol.* **11**: 651-662.

Birnbaum, K., Shasha, D.E., Wang, J.Y., Jung, J.W., Lambert, G.M., Galbraith, D.W., Benfey, P.N. 2003. A gene expression map of the *Arabidopsis* root. *Science.* **302**: 1956-60.

Birnboim, H.C. and Doly, J. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* **7**: 1513-1523.

Blume, B and Grierson, D. 1997. Expression of ACC oxidase promoter-GUS fusions in tomato and *Nicotiana plumbaginifolia* regulated by developmental and environmental stimuli. *Plant J.* **12**: 731-746.

Borisjuk, N.V., Borisjuk, L.G., Logendra, S., Petersen, F., Gleba, Y. and Raskin, I. 1999. Production of recombiant proteins in plant root exudates. *Nat. Biotechnol.* **17**: 466-469

Bouche, N. and Bouchez, D. 2001. *Arabidopsis* gene knockout: Phenotypes wanted. *Curr. Opin. Plant Biol.* **4**: 111-117.

Bouchez, D. and Hofte, H. 1998. Functional genomics in plants. *Plant Physiol* **118**: 725-732.

Breathnach, R. and Chambon, P. 1981. Organization and expression of eukaryotic split genes coding for proteins. *Annu. Rev. Biochem.* **50**: 349-383.

Brenner, R., Thomas, T.O, Becker, M.N. and Atkinson, N.S. 1996. Tissue-specific expression of a Ca2+-activated K+ channel is controlled by multiple upstream regulatory elements. *J. Neuroscience* **16**: 1827-1835.

Butler, J.E.F. and Kadonaga, J.T. 2002. The RNA polymerase II core promoter: a key component in the regulation of gene expression. *Genes Dev.* **16**: 2583-2592.

Caliskan M, Bashiardes S, Ozcan B, Cuming AC. 2003. Isolation and localization of new germination-related sequences from wheat embryos. *J. Biochem. Mol. Biol.* **36**: 580-585.

Cass, L.G., Kirven, K.A. and Christofferson, R.F. 1990. Isolation and characterization of a cellulase gene family member expressed during avocado fruit ripening. *Mol. Gen. Genet.* **223**: 76-86.

Casson, S. A., Chilley, P. M., Topping, J. F., Evans, I. M., Souter, M. A. and Lindsey, K. 2002. The POLARIS gene of *Arabidopsis* encodes a predicted peptide required for correct root growth and leaf vascular patterning. *Plant Cell.* **14**: 1705-1721

Casson, S.A. and Lindsey, K. 2003. Genes and signaling in root development *New Phytologist* **158**: 11-38.

Clemens, S., Palmgren , M.G., and Krämer, U. 2002. A long way ahead: understanding and engineering plant metal accumulation. *Trends in Plant Sci.* **7**: 309-315.

Clough, S. and Bent, A. 1998. Floral dip a simplified method for *Agrobacterium* mediated transformation of *Arabidopsis thaliana. Plant J.* **16**: 735-743.

Crawford, N.M. Nitrate: nutrient and signal for plant growth. 1995. *Plant Cell.* **7**: 859-68.

Cullimore, J.V., Ranjeva, R. and Bono, J.J. 2001. Perception of lipo-chitooligosaccharidic Nod factors in legumes. *Trends Plant Sci* **6**: 24-30.

DeGreve, H., Nguyen, V. K. Deboeck, F., Thia-Toong, L., Karimi, M. Hernalsteens, J. P. 2001. T-DNA tagging of the translation initiation factor eIF-4A1 of *Arabidopsis thaliana. Plant Sci.* **161**: 685-693.

Deikman, J. and Fischer, R.L. 1988. Interaction of a DNA binding factor with the 5' flanking region of an ethylene responsive fruit ripening gene from tomato *EMBO J.* **7**: 3315-3320.

Deikman, J., Xu, R., Kneissel, M.L., Ciardi, J.A., Kim, K.N. and Pelah, D. 1998. Separation of *cis* elements responsive to ethylene, fruit development and ripening in the ripening-related E8 gene. *Plant Mol. Biol.* **37**: 1001-1011.

Doyle, M.C. and Han, I.S. 2001. The roles of two TATA boxes and 3'-flanking region of soybean beta-tubulin gene (tubB1) in light-sensitive expression. *Mol. Cells.* **12**: 197-203.

Drake, P.M., Chargelegue, D.M., Vine, N.D., van Dolleweerd, C.J., Obregon, P. and Ma, J.K. 2003. Rhizosecretion of a monoclonal antibody protein complex from transgenic tobacco roots. *Plant Mol. Biol.* **52**: 233-241.

Elrouby, N. and Bureau, T.E. 2000. Molecular Characterization of the *Abp1* 5′-Flanking Region in Maize and the Teosintes. *Plant Physiol.* **124**: 369-378.

Esau, K. 1977. Anatomy of Seed Plants. *New York: John Wiley and Sons.*

Escobar, C., De Meutter, J., Aristizabal, F.A., Sanz-Alferez, S., del Campo, F.F., Barthels, N., Van der Eycken, W., Seurinck, J., van Montagu, M., Gheysen, G. and Fenoll, C. 1999. Isolation of the LEMMI9 gene and promoter analysis during a compatible plant-nematode interaction. *Mol. Plant Microbe Interact.* **12**: 440-9.

Farrar, K, Evans, I.M, Topping, J. F., Souter, M.A., Nielsen, J. E. and Lindsey, K. 2003. *EXORDIUM*- a gene expressed in proliferating cells and with a role in meristem function, identified by promoter trapping in *Arabidopsis. Plant J.* **33**: 61-73.

Feldmann, K.A. 1991. T-DNA insertion mutagenesis in *Arabidopsis*: mutational spectrum. *Plant J.* **1**: 71-82.

Feldmann, K.A. and Marks, M. D. 1987. *Agrobacterium*-mediated transformation of germinating seeds of *Arabidopsis thaliana*. *Mol. Gen. Genet.* **208**: 1-9.

Feldmann, K.A., Coury, D.A. and Christianson, M.L. 1997. Exceptional segregation of a selectable marker (KanR) in *Arabidopsis* identifies genes important for gametophytic growth and development. *Genetics.* **147**: 1411-1422.

Feng, J.X., Ji, S.J., Shi, Y.H., Xu, Y., Wei, G. and Zhu, Y.X. 2004. Analysis of five differentially expressed gene families in fast elongating cotton fiber. *Acta Biochem. Biophys. Sin (Shanghai).* **36**: 51-56.

Ferrario, S., Immink, R.G., Shchennikova, A., Busscher-Lange, J. and Angenent, G.C. 2003. The MADS box gene FBP2 is required for SEPALLATA function in petunia. *Plant Cell.* **5**: 914-925.

Filleur, S., Dorbe, M.F., Cerezo, M., Orsel, M., Granier, F., Gojon, A. and Daniel-Vedele, F. 2001 An *Arabidopsis* T-DNA mutant affected in Nrt2 genes is impaired in nitrate uptake. *FEBS Lett.* **489**: 220-224.

Fobert, P. R., Labbe, H., Cosmopaulos, J., Gottlob McHugh, S., Quellet, J., Hattori, J., Sunohara, G., Iyer, V. N. and Miki, B. L. 1994. T-DNA tagging of a seed coat specific cryptic promoter in tobacco. *Plant J.* **6**: 567-577.

Forsthoefel, N.R., Wu, Y., Schulz, B., Bennet, M.J., and Feldmann, K.A. 1992. T-DNA insertional mutagenesis in *Arabidopsis*: Prospects and Perspectives. *Aust. J. Plant Physiol.* **19**: 353-356.

Foster, E., Hattori, J., Labbe, H., Ouellet J., Fobert, P. R., James, L. E., Iyer, V. N., Miki, B. L. 1999. A tobacco cryptic constitutive promoter tCUP revealed by T-DNA tagging. *Plant Mol. Biol.* **41**: 45-55.

Franco-Zorrilla, J.M., Gonzalez, E., Bustos, R., Linhares, F., Leyva, A. and Paz-Ares, J. 2004. The transcriptional control of plant responses to phosphate limitation. *J Exp. Bot.* **55**: 285-293.

Fourel, G., Transy, C., Tennant, B.C. and Buendia, M.A. 1992. Expression of the woodchuck N-myc2 retroposon in brain and in liver tumors is driven by a cryptic N-myc promoter. *Mol. Cell Biol.* **12**: 5336-5344.

Gleba, D., Borisjuk, N.V, Borisjuk, L.G., Kneer, R., Poulev, A., Skarzhinskaya, M., Dushenkov, S., Logendra, S., Gleba, Y.Y. and Raskin I. Use of plant roots for phytoremediation and molecular farming. *Proc. Natl. Acad. Sci. USA.* 1999. **96**: 5973-5977.

Goldberg, M.L. 1979. "Sequence analysis of Drosophila histone genes." *Ph.D. dissertation Stanford University, California.*

Goldsbrough, A. and Bevan, M. 1991. New patterns of gene activity in plants detected using an *Agrobacterium* vector. *Plant Mol. Biol.* **16**: 263-269.

Goddijn, O.J., Lindsey, K., van der Lee, F.M., Klap, J.C. and Sijmons, P.C. 1993 Differential gene expression in nematode-induced feeding structures of transgenic plants harbouring promoter-gusA fusion constructs. *Plant J.* **4**:863-873.

Grace, M.L., Chandrasekharan, M.B., Hall, T.C. and Crowe, A.J. 2004. Sequence and spacing of TATA box elements are critical for accurate initiation from the beta-phaseolin promoter. *J Biol Chem.* **279** :8102-8110.

Hanahan, D. 1983. Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* **166**: 557-580.

Hansen, E., Harper, G., McPherson, M.J. and Atkinson, H.J. 1996. Differential expression of patterns of the wound-inducible transgene *wun1-uidA* in potato roots following infection with either cyst or root knot nematodes. *Physiol. Mol. Plant Pathol.* **48**: 161–170.

Haran, S., Logendra, S., Seskar, M., Bratanova, M. and Raskin, I. 2000. Characterization of *Arabidopsis* acid phosphatase promoter and regulation of acid phosphatase expression. *Plant Physiol.* **124**: 615-626.

Hattori, T., Nakagawa, S. and Nakamura, K. High-level expression of tuberous root storage protein genes of sweet potato in stems of plantlets grown in vitro on sucrose medium. *Plant Mol. Biol.* 1990 **14**: 595-604.

Hayashi, H., Czaja, I., Lubenow, H., Schell, J. and Walden, R. 1992. Activation of a plant gene by T-DNA tagging: auxin-independent growth in vitro. *Science.* **258**: 1350-1353.

Herman, P.L. and Marks, M.D. 1989. Trichome Development in *Arabidopsis thaliana*. II. Isolation and Complementation of the GLABROUS1 Gene. Plant Cell. **1**: 1051-1055.

Herman, L., Jacobs, A., Van Montagu, M. and Depicker, A. 1990. Plant chromosome/marker gene fusion assay for study of normal and truncated T-DNA insertion events. *Mol. Gen. Genet.* **224**: 248-256.

Hild, M., Beckmann, B., Haas, S.A., Koch, B., Solovyev, V., Busold, C., Fellenberg, K., Boutros, M., Vingron, M., Sauer, F., Hoheisel, J.D. and Paro, R. 2003. An integrated gene annotation and transcriptional profiling approach towards the full gene content of the Drosophila genome. *Genome Biol.* **5**: R3.1-17.

Hirsch, E.R., Lewis, B.D., Spalding, E.P. and Sussman, M.R. 1998. A role for the AKT1 Potassium channel in plant nutrition. *Science.* **280**: 918-921.

Hirsch, A.M., Lum, M.R. and Downie, J.A. 2001. What makes the rhizobia-legume symbiosis so special? *Plant Physiol.* **127**: 1484-1492.

Ho, M.W., Ryan, A. and Cummins, J. Cauliflower Mosaic Viral Promoter – A Recipe for Disaster? 1999. *Microbial Ecology in Health and Disease*, **11**: 194-197.

Hohnjec, N., Perlick, A.M., Puhler, A. and Kuster, H. 2003. The Medicago truncatula sucrose synthase gene MtSucS1 is activated both in the infected region of root nodules and in the cortex of roots colonized by arbuscular mycorrhizal fungi. *Mol. Plant Microbe Interact.* **16**: 903-915.

Honma, T. and Goto, K. 2001. Complexes of MADS-box proteins are sufficient to convert leaves into floral organs. *Nature.* **409**: 525-529.

Hua, X., van de Cotte, B., Van Montagu, M. and Verbruggen, N. 1999. A 69 bp fragment in the pyrroline-5-carboxylate reductase promoter of *Arabidopsis thaliana* activates minimal CaMV 35S promoter in a tissue-specific manner *FEBS Lett.* **458**: 193-196.

Huang, S., Cerny, R.E., Bhat, D.S. and Brown, S.M. 2001. Cloning of an *Arabidopsis* patatin-like gene, STURDY, by activation T-DNA tagging. *Plant Physiol.* **129**:1544-1556.

Irniger, S., Egli, C.M., Kuenzier, M. and Brauss, G.H. 1992. The yeast actin intron contains a cryptic promoter that can be switched on by preventing transcriptional interference. *Nucleic Acids Res.* **20**: 4733-4739.

Jach, G., Gornhardt, B., Mundy, J., Logemann, J., Pinsdorf, E., Leah, R., Schell, J. and Mass, C. 1995. Enhanced quantitative resistance against fungal diseases by combinatorial expression of different barley antifungal proteins in transgenic tobacco. *Plant J.* **8**: 97-109.

Jefferson, R. A., Kavanagh, T. A. and Bevan, M. W. 1987. GUS fusion: β-Glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J.* **6**: 3901-3907.

Jeong, D. H., An, S., Kang, H. G., Moon, S., Han, J. J., Park, S., Lee, H. S., An, K. and An, G. 2002. T-DNA insertional Mutagenesis for activation tagging in rice. *Plant Physiol.* **130**:1636-1644.

Kardailsky, I., Shukla, V.K., Ahn, J.H., Dagenais, N., Christensen, S.K., Nguyen, J.T., Chory, J., Harrison, M.J. and Weigel, D. 1999. Activation tagging of the floral inducer FT. *Science.* **286**:1962-1965.

Kertbundit, S., De Greve, H., Deboeck, F., Van Montagu, M. and Hernalsteens, J.-P. 1991. *In vivo* random 2-glucuronidase gene fusions in *Arabidopsis thaliana* . *Proc. Natl. Acad. Sci.* USA, **88**: 5212-5216.

Khan, A.G., Kuek, C., Chaudhry, T.M., Khoo, C.S. and Hayes, W.J. 2000. Role of plants, mycorrhizae and phytochelators in heavy metal contaminated land remediation. *Chemosphere*. **41**: 197-207.

Kim, E.J., Kwak, J.M., Uozumi, N. and Schroeder, J.I. 1998. AtKUP1: an *Arabidopsis* gene encoding high-affinity potassium transport activity. *Plant Cell*. **10**: 51-62.

Kleckner, N., Chan, R.K., Tye, B.K. and Botstein, D. 1975. Mutagenesis by insertion of a drug-resistance element carrying an inverted repetition. J Mol Biol. **97**: 561-575.

Koltunow, A. M., Truettner, J., Cox, K. H., Wallroth, M. and Goldberg, R. B. 1990. Different temporal and spatial gene expression patterns occur during anther development. *Plant Cell*. **2**: 1201-1224.

Koncz, C. and Schell, J. 1986. The promoter of the T_L DNA gene 5 controls the tissue specific expression of chimeric genes carried by a noval type *Agrobacterium* binary vector. *Mol. Gen. Genet.* **204***:* 383-396.

Koncz, C., Martini, N., Mayerhofer, R., Koncz-Kalman, Z., Korber, H., Redei, G.P. and Schell, J. 1989. High frequency T-DNA gene tagging in plants. *Proc. Natl. Acad. Sci. USA*. **86**: 8467-8471.

Koncz, C., Mayerhofer, R., Koncz-Kalman, Z., Nawrath, C., Reiss, B., Redei, G.P. and Schell, J. 1990. Isolation of a gene encoding a novel chloroplast protein by T-DNA tagging in *Arabidopsis thaliana. EMBO J.* **9**:1337-46.

Krysan,J.P., Young, J.C., Tax, F. and Sussman, M.R. 1996. Identification of transferred DNA insertions within *Arabidopsis* genes involved in signal transduction and ion transport. *Proc. Natl. Acad. Sci. USA*. **93**: 8145-8150.

Kuiper, I., Lagendijk, E.L., Bloemberg, G.V. and Lugtenberg, B.J.J. 2004. Rhizoremediation: A Beneficial Plant-Microbe Interaction. *Mol. Plant-Microbe Interact*. **17**: 6-15.

Kutach, A.K. and Kadonaga, J.T. 2000. The downstream promoter element DPE appears to be as widely used as the TATA box in *Drosophila* core promoters. *Mol. Cell. Biol.* **20**: 4754-4764.

Lagarde, D., Basset, M., Lepetit, M., Conejero, G., Gaymard, F., Astruc, S. and Grignon, C. 1996. Tissue-specific expression of *Arabidopsis* AKT1 gene is consistent with a role in K+ nutrition. *Plant J.* **9**:195-203.

Li,Q. and Nicholas,R.A. 1997. A PCR strategy to obtain 5' and 3' flanking sequences of partial clones from lambda Zap c DNA libraries of Partial Clones from lambdaZAP cDNA Libraries, *BioTechniques.* **22**:1038-1044.

Lilley, C.J., Urwin, P.E. and Atkinson, H.J. 1999. Characterization of plant nematode genes: identifying targets for a transgenic defence.*Parasitology.***118**. Suppl: S63-72.

Lindsey, K., Wei, W., Clarke, M. C., McArdle, H. F., Rooke, L. M. and Topping, J. F. 1993. Tagging genomic sequences that direct transfer expression by activation of a promoter trap in plants. *Transgenic Res.* **2**: 33-47.

Lindsey K, Topping JF, Muskett PR, Wei W, Horne KL. 1998. Dissecting embryonic and seedling morphogenesis in *Arabidopsis* by promoter trap insertional mutagenesis. *Symp. Soc. Exp. Biol.* **51**:1-10.

Liu, Y., Mitsukawa, N., Oosumi, T. Whittier, R. 1995. Efficient isolation and mapping of *Arabidopsis thaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR. *Plant J.* **8**: 457-463.

Liu, H., Trieu, A.T, Blaylock, L.A. and Harrison, M.J.1998. Cloning and characterization of two phosphate transporters from Medicago truncatula roots: regulation in response to phosphate and to colonization by arbuscular mycorrhizal (AM) fungi. *Mol. Plant Microbe Interact.* **11**:14-22.

Lugtenberg, B.J.J. and Dekkers, L.C. 1999. What makes *Pseudomonas* bacteria rhizosphere competent. *Environ. Microbiol.* **1**: 9-13.

Lugtenberg, B.J.J., Dekkers, L.C. and Bloemberg, G.V. 2001. Molecular determinants of rhizosphere colonization by *Pseudomonas. Annu. Rev. Phytopathol.* **39**: 461-490.

Mandal, A., Sandgen, M., Holmstrom, K. O., Gallois, P., Palua, E. T. 1995. Identification of *Arabidopsis thaliana* sequences responsive to low temperature and abscissic acid by T-DNA tagging and *in vivo* gene fusion. *Plant Mol. Biol. Rep.* **13**: 243-254.

Marks, M.D. and Feldmann, K.A. 1989. Trichome Development in *Arabidopsis thaliana*. I. T-DNA Tagging of the GLABROUS1 Gene. Plant Cell. **1**: 1043-1050.

Martienssen, R.A, Barkan, A., Freeling, M. and Taylor, W.C. 1989. Molecular cloning of a maize gene involved in photosynthetic membrane organization that is regulated by Robertson's Mutator. *EMBO J.* **8**:1633-9.

Matsuyama, T., Satoh, H., Yamada, Y. and Hashimoto, T. 1999. A maize glycine-rich protein is synthesized in the lateral root cap and accumulates in the mucilage. *Plant Physiol.* **120**: 665-674.

Meagher, R.B. 2000. Phytoremediation of toxic elemental organic pollutants. *Curr. Opin. Plant Biol.* **3**:162.

Meinke, D.W., Meinke, L.K., Showalter, T.C., Schissel, A.M., Mueller, L.A. and Tzafrir, I. 2003. A sequence based map of *Arabidopsis* genes with mutant phenotypes. *Plant Physiol.* **131**. 409-418.

Meissner, R.C., Jin, H., Cominelli, E., Denekamp, M., Fuertes, A., Greco, R., Kranz, H.D., Penfield, S., Petroni, K., Urzainqui, A., Martin, C., Paz-Ares, J., Smeekens, S., Tonelli, C., Weisshaar, B., Baumann, E., Klimyuk, V., Marillonnet, S., Patel, K., Speulman, E., Tissier, A.F., Bouchez, D., Jones, J.J., Pereira, A., Wisman, E. and Bevan, M. 1999. Function search in a large transcription factor gene family in *Arabidopsis*: assessing the potential of reverse genetics to identify insertional mutations in R2R3 MYB genes. *Plant Cell.* **11**:1827-40.

Meister, R.J., Williams, L.A., Monfared, M.M., Gallagher, T.L., Kraft, E.A., Nelson, C.G., Gasser CS. 2004. Definition and interactions of a positive regulatory element of the *Arabidopsis* INNER-NO-OUTER promoter *Plant J.* **37**:426-438.

Mitra, R.M., Gleason, C.A., Edwards, A., Hadfield, J., Downie, J.A., Oldroyd, G.E and Long, S.R. A Ca2+/calmodulin-dependent protein kinase required for symbiotic nodule development: Gene identification by transcript-based cloning. *Proc. Natl. Acad. Sci. U S A.* **101**: 4701-4705.

Mollier, P., Montoro, P., Delarue, M., Bechtold, N., Bellini, C. and Pelletier, G. 1995. Promoter less gusA expression in a large number of *Arabidopsis thaliana* transformants obtained by *in planta* infiltration method. *C. R. Acad. Sci.* Paris, **318**: 465-474.

Mollier, P., Hoffman, B., Orsel, M., and Pelletier, G. 2000. Tagging of a cryptic promoter that confers root-specific *gus* expression in *Arabidopsis thaliana*. *Plant Cell Rep.* **19**: 1076-1083.

Montgomery, J., Goldman, S., Deikman, J., Margossian, L, Fischer, R.L. 1993a. Identification of an ethylene responsive region in the promoter of a fruit-ripening gene. *Proc. Natl. Acad. Sci. USA.* **90**: 5939-5943.

Montgomery, J., Pollard, V., Deikman, J., Fischer, R.L.1993b. Positive and negative regulatory regions control the spatial distribution of polygalacturonase transcription in tomato fruit pericarp. *Plant Cell.* **5**: 1049-1062.

Nakamura, M., Tsunoda, T. and Obokata, J. 2002. Photosynthesis nuclear genes generally lack TATA-boxes: a tobacco photosystem I gene responds to light through an initiator. *Plant J.* **29**: 1-10.

Nakazawa, M., Yabe, N., Ichikawa, T., Yamamoto, Y.Y., Yoshizumi, T., Hasunuma, K. and Matsui, M. 2001. DFL1, an auxin-responsive GH3 gene homologue, negatively regulates shoot cell elongation and lateral root formation, and positively regulates the light response of hypocotyl length. *Plant J.* **25**: 213-221.

Narvaez-Vasquez, J. and Ryan, C.A. 2002. The systemin precursor gene regulates both defensive and developmental genes in Solanum tuberosum. *Proc Natl Acad Sci U S A.* **99**: 15818-15821.

Nazoa, P., Vidmar, J.J., Tranbarger, T.J., Mouline, K., Damiani, I., Tillard, P., Zhuo, D, Glass, A.D. and Touraine, B. 2003. Regulation of the nitrate transporter gene AtNRT2.1 in *Arabidopsis thaliana*: responses to nitrate, amino acids and developmental stage. *Plant Mol Biol.* **52**: 689-703.

Ohkama, N., Goto, D.B., Fujiwara, T., Naito, S. 2002. Differential tissue-specific response to sulfate and methionine of a soybean seed storage protein promoter region in transgenic *Arabidopsis*. *Plant Cell Physiol.* **43**:1266-1275.

Okada, K. and Shimura, Y. 1990. Reversible root tip rotation in *Arabidopsis* seedlings induced by obstacle touching stimulus. *Science.* **250**: 274-276.

Ökrész, L., Máthé, C., Horváth, E., Schell, J., Koncz, C. and Szabados , L. 1998. T-DNA trapping of a cryptic promoter identifies an ortholog of highly conserved *SNZ* growth arrest response genes in *Arabidopsis*. *Plant Sci.* **138**: 217-228.

Orsel, M, Filleur, S., Fraisier, V., Daniel-Vedele, F. 2002. Nitrate transport in plants: which gene and which control? *J. Exp. Bot.* **53**: 825-833.

Parinov, S., Sevugan, M., De, Y., Yang, W. C., Kumaran, M. and Sundaresan, V. 1999. Analysis of flanking sequences from dissociation insertion lines: A database for reverse genetics in *Arabidopsis*. *Plant Cell.***11:** 2263-2270.

Pelaz, S., Tapia-Lopez, R., Alvarez-Buylla, E.R and Yanofsky, M.F. 2001 Conversion of leaves into petals in *Arabidopsis*. *Curr Biol.* **11**:182-184.

Plesch, G., Kamann, E. and Mueller, R. B. 2000. Cloning of regulatory sequences mediating guard-cell-specific gene expression. *Gene,* **249**: 83-89.

Puzio, P. S., Lausen, A., Almmeida- Engler, J., Cai, D., Gheysen, G., and Grundler, F. M. 1999. Isolation of a gene from *Arabidopsis thaliana* related to nematode feeding structures. *Gene.* **239**: 163-172.

Qin, G., Kang, D., Dong, Y., Shen, Y., Zhang, L., Deng, X., Zhang, Y., Li, S., Chen, N., Niu, W., Chen, C., Liu, P., Chen, H., Li, J., Ren, Y., Gu, H., Deng, X.W., Qu, L.J. and Chen, Z. 2003. Obtaining and analysis of flanking sequences from T-DNA transformants of *Arabidopsis*. *Plant Sci.* **165**: 941-949.

Quinter, F.J. and Blatt, M.R. 1997. A new family of K+ transporters from *Arabidopsis* that are conserved across phyla. *FEBS Lett.* **415**: 206-211.

Resminath, R. 2003. Promoter tagging in *Arabidopsis* using T-DNA insertional mutagenesis. *PhD Thesis. IARI, New Delhi.*

Richardson, K., Fowler, S., Pullen, C., Skelton, C., Morris, B. and Putterill, J., 1998. T-DNA tagging of a flowering-time gene and improved gene tranfer by *in planta* transformation of *Arabidopsis*. *Aust. J. Plant Physiol.* **25**: 125-130.

Roa-Rodriguez, C. 2003. Promoters used to regulate gene expression. *The Team, CAMBIA intellectual Property Resource.* Canberra, Australia.

Ross, K.J., Fransz, P., Armstrong, S.J., Vizir, I., Mulligan, B., Franklin, F.C. and Jones, G.H. 1997. Cytological characterization of four meiotic mutants of *Arabidopsis* isolated from T-DNA-transformed lines. *Chromosome Res.* **5**: 551-559.

Rosso, M.G., Li, Y., Strizhov, N., Reiss, B., Dekker, K. and Weisshaar, B. 2003. An *Arabidopsis thaliana* T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics. *Plant Mol Biol.* **53**: 247-259.

Saedler, H., Becker, A., Winter, K.U., Kirchner, C. and Theissen, G. MADS-box genes are involved in floral development and evolution. 2001. *Acta Biochem Pol.* **48**: 351-358.

Sambrook, J., Fritsch, E.F. and Maniatis, T. 1989. Molecular Cloning - A Laboratory Manual, 2nd Edition. *Cold Spring Harbor Laboratory Press, New York.*

Santino, C.G., Stanford, G.L. and Conner, T.W. 1997. Developmental and transgenic analysis of two tomato fruit enhanced genes. *Plant Mol. Biol.* **33**: 405-416.

Schiefelbein, J.W. and Benfey, P.N. 1991. The development of plant roots: new approaches to underground problems. *Plant Cell.* **2**: 235-243

Schindelman, G., Morikami, A., Jung, J., Baskin, T.I., Carpita, N.C., Derbyshire, P., McCann, M.C. and Benfey, P.N. 2001. COBRA encodes a putative GPI-anchored protein, which is polarly localized and necessary for oriented cell expansion in *Arabidopsis. Genes Dev.***15:** 1115-1127.

Singer, V.L., Wobbe, C.R., and Struhl, K. 1990. A wide variety of DNA sequences can functionally replace a yeast TATA element for transcriptional activation. *Genes Dev.* **4**: 636-645.

Singh, K.B. 1998. Transcriptional Regulation in Plants: The importance of Combinatorial Control. *Plant Physiol.* **118**: 1111-1120.

Speulman, E., Metz, P. L., van Arkel G., te Lintel Hekkert, B., Stiekema, W. J. and Pereira, A. 1999. A two-component *Enhancer-inhibitor* transposon mutagenesis system for the functional analysis of the *Arabidopsis* genome. *Plant Cell.* **11:** 1853-1866.

Sjodahl, S., Gustavsson, H.O, Rodin, J. and Rask, L. 1995. Deletion analysis of the *Brassica napus* cruciferin gene cru 1 promoter in transformed tobacco: promoter activity during early and late stages of embryogenesis is influenced by cis-acting elements in partially separate regions. *Planta*: **197**: 264-271.

Smale, S.T. 2001. Core promoters: Active contributors to combinatorial gene regulation. *Genes Dev.* **15**: 2503-2508.

Sommer, H., Beltran, J.P., Huijser, P., Pape, H., Lonnig, W.E., Saedler, H. and Schwarz-Sommer, Z. 1990. Deficiens, a homeotic gene involved in the control of flower morphogenesis in Antirrhinum majus: the protein shows homology to transcription factors. EMBO J. **9**:605-13.

Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**: 503-517.

Sriprang, R., Hayashi, M., Yamashita, M., Ono, H., Saeki, K. and Murooka, Y. A novel bioremediation system for heavy metals using the symbiosis between leguminous plant and genetically engineered rhizobia. *J Biotechnol.* 2002. **99**: 279-293.

Steeves, T.A. and Sussex, I.M. 1989. Patterns in Plant Development. *Cambridge: Cambridge University Press.*

Stephen, T. and Smale, 2001. Core promoters: active contributors to combinatorial gene regulation. *Genes Dev.* **15**: 2503-2508.

Sundaresan, V., Springer, P., Volpe, T., Haward, S., Jones,J. D. G., Dean, C, Ma, H. and Martienssen, R. 1995. Patterns of gene action in plant development revealed by enhanced trap and gene trap transposable elements. *Genes Dev.* **9**: 1797-1810.

Sunilkumar, G., Connell, J.P., Smith, C.W., Reddy, A.S. and Rathore, K.S. 2002. Cotton alpha-globulin promoter: isolation and functional characterization in transgenic cotton, *Arabidopsis*, and tobacco. *Transgenic Res.***11**: 347-359.

Suzuki, Y., Tsunoda, T., Sese, J., Taira, H., Mizushima-Sugano, J., Hata, H., Ota, T., Isogai, T., Tanaka, T., Nakamura, Y. et al. 2001. Identification and characterization of the potential promoter regions of 1031 kinds of human genes. *Genome Res.* **11**: 677-684.

Takahashi, N., Hotta, H. and Homma, M., 1991. Activation and suppression of a cryptic promoter in the intron of the human melanoma-associated ME491 antigen gene. *Jpn J. Cancer Res.* **82**: 1239-1244.

Topping, J. F., Agyeman, F., Henricot, B. and Lindsey, K. 1994. Functional tagging of regulatory elements in the plant genome. *Development.* **112**: 1009-1019.

Topping, J.F. and Lindsey, K. 1997. Promoter trap markers differentiate structural and positional components of polar development in *Arabidopsis. Plant Cell.* **9**:1713-1725.

Tissier, A. F., Marillonnet, S., Klimyuk, V., Patel, K., Torres, M. A., Murphy, G., Jones, J. D. G. 1999. Multiple independent defective *Suppressor-Mutator* transposon insertions in *Arabidopsis*: a tool for functional genomics. *Plant Cell*, **11**: 1841-1852.

Tsugeki, R. and Fedoroff, N.V. 1999. Genetic ablation of root cap cells in *Arabidopsis. Proc. Natl. Acad. Sci. U S A.* **96**:12941-12946.

Tsunoda, T. and Takagi, T. 1999. Estimating transcription factor bindability on DNA. *Bioinformatics.* **15**: 622-630.

Tzafrir, I., Dickerman, A., Brazhnik, O., Nguyen, Q., McElver, J., Frye, C., Patton, D. and Meinke, D. 2003. The Arabidopsis SeedGenes Project. *Nucleic Acids Res.* **31**: 90-93.

Ukai, H., Ukai-Tadenuma, M., Ogiu, T and Tsuji, H. 2002. A new technique to prevent self-ligation of DNA. *J. Biotechnol.* **97**: 233-242.

Vaghchhipawala, Z. E. Schlueter, J.A. Shoemaker, R.C. and Mackenzie, S.A. Soybean FGAM synthase promoters direct ectopic nematode feeding site activity. 2004. *Genome.* **47**: 404-413.

Valvekens, D., Van Montagu, M. and Van Lijse bettens, M. 1988. *A. tumefaciencs* mediated transformation of *Arabidopsis thaliana* root explants by using kanamycin selection. *Proc. Natl. Acad. Sci.* USA **85**: 5536-5546.

van der Graaff, E., Dulk-Ras, A.D., Hooykaas, P.J. and Keller, B. 2000. Activation tagging of the LEAFY PETIOLE gene affects leaf petiole development in *Arabidopsis thaliana. Development.* **127**: 4971-80.

Van Haaren, M.J.J. and Houck, C.M. 1991. Strong negative and positive regulatory elements contribute to the high-level fruit specific expression of the tomato 2AII gene. *Plant Mol. Biol.* **17**: 615-630.

Van Haaren, M.J.J. and Houck, C.M.1993. A functional map of the fruit-specific promoter of the tomato 2AII gene. *Plant Mol. Biol.* **21**: 625-640.

Vander Mijnsbrugge, K., Van Montagu, M, Inze, D. and Boerjan, W. 1996. Tissue-specific expression conferred by the S-adenosyl-L-methionine synthetase promoter of *Arabidopsis thaliana* in transgenic poplar. *Plant Cell Physiol.* **37**: 1108-15.

Vieweg, M.F., Fruhling, M., Quandt, H.J., Heim, U., Baumlein, H., Puhler, A., Kuster, H. and Andreas, M.P. 2004. The promoter of the Vicia faba L. leghemoglobin gene VfLb29 is specifically activated in the infected cells of root nodules and in the arbuscule-containing cells of mycorrhizal roots from different legume and nonlegume plants. *Mol Plant Microbe Interact.* **17**: 62-69.

Wang, Y, Wang, X., McCubbin, A.G. and Kao, T.H. 2003. Genetic mapping and molecular characterization of the self-incompatibility (S) locus in Petunia inflata. *Plant Mol Biol.* **53**: 565-580.

Webb, K.J., Skot, L., Nicholson, M.N., Jorgensen, B. and Mizen, S. 2000 *Mesorhizobium loti* increases root-specific expression of a calcium-binding protein homologue identified by promoter tagging in *Lotus japonicus. Mol Plant Microbe Interact.* **13**: 606-616.

Wei, W., Twell, D. and Lindsey, K. 1997. A novel nucleic acid helicase gene identified by promoter trapping in *Arabidopsis*. *Plant J.* **11**: 1307-1314.

Weigel, D., Ann, J. H., Blazquez. M. A., Borevitz, J. O., Christenoen, S. K., Fankhauser, C., Ferrandiz, C., Kardialsky, I., Malanchosuil, E. J. and Neff, M. M. 2000. Activation tagging in *Arabidopsis*. *Plant Physiol.* **122**: 1003-1013.

Weigel, D. and Glazebrook, J. 2002. Arabidopsis a laboratory manual, Cold Spring Harbor Laboratory Press.

Willemsen, V., Wolkenfelt, H., de Vrieze, G., Weisbeek, P. and Scheres, B. 1998). The *HOBBIT* gene is required for formation of the primary root meristem in the *Arabidopsis* embryo. *Development.* **125**: 521-531.

Wu, K., Malik, K., Tian, L., Hu, M., Martin, T., Foster, L., Brown, D. and Miki, B. 2001. Enhancers and core promoter elements are essential for the activity of a cryptic gene activation sequence. *Mol. Genet. Genomics.* **265**: 763-770.

Wu, K., Hu, M., Martin, T., Wang, C., Li, X., Tian, L., Brown, D. and Miki, B. 2003. The cryptic enhancer elements of the *tCUP* promoter. *Plant Mol. Biol.* **51**: 351-362.

Xiao, K, Wang, Z. and Harrison, M, 2003. Isolation and analysis of two root specific promoters from *Medicago truncatula*. *Third international Symposium on Molecular Breeding of Forage and Turf, May 18-22, 2003, Dallas Texas, USA.*

Xu, R., Goldman, S., Coupe, S. and Deikman, J. 1996. Ethylene control of E4 transcription during tomato fruit ripening involves two co-operative *cis* elements. *Plant Mol. Biol.* **31**: 1117-1127.

Yamaguchi-Shinozaki, K. and Shinozaki, K. 1994. A novel cis-acting element in an Arabidopsis gene is involved in responsiveness to drought, low-temperature, or high-salt stress. *Plant Cell.* **6**: 251-264.

Yamamoto, S, Nishihara, M., Morikawa, H., Yamauchi, D. and Minamikawa, T. 1995. Promoter analysis of seed storage protein genes from *Canavalia gladiata* D.C. *Plant Mol Biol.* **27**: 729-41.

Yanofsky, M.F., Ma, H., Bowman, J.L., Drews, G.N., Feldmann, K.A. and Meyerowitz, E.M. 1990.The protein encoded by the *Arabidopsis* homeotic gene agamous resembles transcription factors. *Nature.* **346**: 35-39.

Yanzhang, W., Guanqiao, Y.U., Shanjiong, S., and Jiabi, Z. 2004. Promoter of soybean early nodulin gene *enod2B* is induced by rhizobial Nod factors in transgenic rice. *Chinese Science Bull.* **49**: 343-349.

Yoshihara, T., Washida, H. and Takaiwa, F. 1996. A 45-bp proximal region containing AACA and GCN4 motif is sufficient to confer endosperm-specific expression of the rice storage protein glutelin gene, GluA-3. *FEBS Lett.* **3833**:213-218.

Zhang, H. and Forde, B.G. 1998. An *Arabidopsis* MADS box gene that controls nutrient-induced changes in root architecture. *Science.* **279**: 407-409.

Zhang, P., Bohl-Zenger, S., Pounti-Kaerlas, J., Potrykus, I. and Gruissem, W. 2003. Two cassava promoters related to vascular expression and storage root formation. *Planta.* **218**: 192-203.

APPENDICES

A.1: Bacterial Growth media and Plasmid extraction buffers:

1. LB medium (pH 7.0)

Chemical	g/l
Tryptone	10
Yeast Exctract	5
NaCl	10

2. LA Agar (pH 7.0)

Chemical	g/l
Tryptone	10
Yeast Exctract	5
NaCl	10
Agar	15

3. YEM Medium

Chemical	g/l
Yeast Extract	0.4
Mannitol	10
NaCl	0.1
$MgSO_4.7H_2O$	0.2
K_2HPO_4	0.5

4. Composition of Plasmid Extraction buffers

Buffer P1 (Resuspension Buffer)
50 mM Tris-Cl, pH 8.0
10 mM EDTA
100 µg/ml RNase A

Buffer P2 (Lysis Buffer)
200 mM NaOH
1% SDS

Buffer P3 (Neutralization Buffer)
3.0 M potassium acetate
pH 5.5

A.2: Arabidopsis Growth Media

1. Murashige and Skoog medium (MS) modified, pH 5.80

Micro elements	(mg/l)
$CoCl_2.6H_2O$	0.025
$CuSO_4.5H_2O$	0.025
FeNaEDTA	36.70
H_3BO_3	6.20
KI	0.83
$MnSO_4.H_2O$	16.90
$Na_2MoO_4.2H_2O$	0.25
$ZnSO_4.7H_2O$	8.60

Macro elements	
$CaCl_2$	332.02
KH_2PO_4	170.0
KNO_3	1900.0
$MgSO_4$	180.54
NH_4NO_3	1650.0

Vitamins	
Myo-inositol	100.0
Nicotinic acid	0.50
Pyridoxine HCl	0.50
Thiamine HCl	0.10
MES	500.0
Sucrose	30,000

2. Hoagland's solution

KNO_3	5 mM
KH_2PO_4 (pH 6.5)	2.5 mM
$MgSO_4$	2.0 mM
$Ca(NO_3)_2$	2.0 mM
Fe-EDTA	50 μM

Micronutrient stock* 1ml/Lit

Adjust pH 6.5 with 2 M KOH

*Micronutrient Stock (70 mM H_3BO_3, 14 mM $MnCl_2$, 0.5 mM $CuSO_4$, 1mM $ZnSO_4$, 0.2 mM Na_2MoO_4, 10 mM NaCl and 0.01 mM $CoCl_2$)

A.3 Gus Assay Buffer

Phosphate Buffer (pH 7.0)	50 mM
EDTA	10 mM
Triton X-100	0.1 %
Potassium Ferrocyanide	1 mM
Potassium Ferricyanide	1 mM
X-Gluc. (Amresco, USA)	1 mM
Methanol	20 %

A.4 Plant DNA Extraction: CTAB Extraction buffer

Stock solution of Extraction Buffer

NaCl	5M
Tris HCl (pH 8.0)	1.0M
Na_2 EDTA (pH 8.0)	0.5M
CTAB	4%
(Dissolved in water overnight)	
β-Mercaptoethanol (ME)	98%
(Added just before extraction)	
Sterile distilled water (SDW)	

A.5 Electrophoresis buffer:

Tris-acetate(TAE)	
50x:	per l
Tris	242g
Glacial Acetic Acid	57.1 ml
0.5 M EDTA (pH 8.0)	40 ml
Working solution is 1x	

A.6 Gel Loading Buffer

6X	
Bromophenol Blue	0.25%
Xylene Cyanol FF	0.25%
Sucrose in water	40% w/v

A.7 Southern Hybridization

1. *Modified Church and Gilbert prehybridization buffer*

Phosphate buffer (pH 7.2)	0.5 M
SDS	7 % (w/v)
EDTA	10 mM

2. 20X SSC

NaCl	3 M
Na_3 citrate	0.3 M

3. Denaturation solution

NaCl	1.5 M
NaOH	0.5 M

4. Neutralization solution

NaCl	1.5 M
Tris-Cl (pH 7.2)	0.5 M
EDTA	1 mM

A.8 Sequences of primers used

1. *gus*-specific

gus Fwd	5' GCC ATT TGA AGC CGA TGT CAC GCC-3'
gus Rev	5'GTA TCG GTG TGA GCG TCG CAG AAC 3'

2. Inverse PCR

TR3	5' GGG TTT CTA CAG GAC GTA AC 3'
LB2	5' CCA CGG AAT AGT TTT GGC CAG ACC 3'

3. Cloning of upstream regulatory sequences
(The magenta colored bases refer to the *Bam*H I site while those colored green refer to *Sal* I site)

Upstream of T-DNA insertion:

R1	5' CTG TTT GGA TCC GGA CCT GCG CGT3'
R2	5' GTC GAC GCT ATG TAT CAC CCG GAT G 3'
R3	5'ATG TCG ACA TAT CGG GTT GAG TCG 3'

Upstream of SAHH gene:

AHB	5' GGA TCC GCT AGA TCT GAG ATT TG 3'
AHB0.5	5' GTC GAC TGA GTG ATA CCA TTA GCG CG 3'
AHB1	5' AGT CGA CTC CAC CGA CGT TCC T 3'

A.9: Sequence of the 1.5 kb inverse PCR Product

Right Border flanking Sequence

pGEMT Sequences ← TR3 →

TGGTCGACCTGCAGGCGGCCGCGAATTCACTAGTGATTGGGTTTCTACAGGACGTAACATAAGGGACTGACCTACCCGGG

← pCS1 Sequences →

GATCCACTAGTTCTAGCCAAGTGCAGCGTTTAATAAATCGTGTGTGTGTGCTGTTTGGATCTGGACCTGCGCGTGTAG

GAACGTCGGTGGATCTGACTGATCTTAAAAGATCAAAACAGAATTAAACAATGTGGTATTAATTATATTAATTATACTTT

ACTATATATATATTTATTGTAATTATATTATGTTGTGTGATTTACCAAAAGAATCGTAATATTTGAAAGCAGCGGACAGT

GAGAATCATAGCAATTATGTGCGATGTGCGTTTCTTTTTTCTTGTTTTGGTGGTACTATGCAGATCTAACATGTTACTA

TCTTTATATGTTTATGAATTTTAAATGTACTTTTCGAATTTTAAATAGATCTTGTTCTTCGATAGAAACTATATACATTGT

ACTACTTTTGCTATTAAACGACTAACAATCGTATATTTCAATAAATTTATAATAGTCTATCTCCACGCATAATGTTTCAC

ATCCGGGTGATACATAGCACTGACC

Left Border Flanking Sequence

◀ PGEM T ▶

GGGAATTCGATTCCACGGAATAGTTTTGGCCAGACCCTTGAAAATCCGATTCAGTACAATCGATTGCCCTCATTTTTTCG

TTAGCATATTGATACTCTTTCTTCAAAAACAAGTTATTGTTTTTATTTTCAACCCAACTTTAATACGGATTCATACTGGG

ATTTAGGTGTTAAATCTGATAATTTAGGTTTGAATAAGTTGTATATTTGTTTCTTTGATTAAAAAAAGAACCTATATATA

TACGAAAATAAATAAAAAGTTCTAGATTTCAATTTTCCGTATATAGCGGGTTGAATTGTCTATTTTAATATGAAAATTGA

CGGATCTTATAAACAAAATGTTCTGAAATATGTAAAAGGATTTAGCCAAAGTTAACCAAAAAAAAAAAAAACAAACAGAAA

AGTCACATTCACATGTCGTGGTAGATCTAAGGCATTAATTTAGAAATATGTCGTTACAATAAGCGGAGAACATGGGACGT

TTCTCGTGGTCCAATCAGACGAACGAGATCTCATAAATTAAATGACTTCAGACGAGGAATTCATGGCAGAATGATAATG

CACCTTAAGTGACTTTAGAGTGAAAATGATACGAGAACAATGCATAATCCA

A.10: Sequence of the cryptic promoter submitted to the Genbank database Accn. No. AY601849

```
AY601849. Arabidopsis thali...[gi:47174865]
LOCUS       AY601849                 446 bp    DNA
linear   PLN 19-MAY-2004
DEFINITION  Arabidopsis thaliana cryptic promoter
region.
ACCESSION   AY601849
VERSION     AY601849.1  GI:47174865
KEYWORDS    .
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta;
Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta;
eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales;
Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 446)
  AUTHORS   Sivanandan,C., Prasad,A.M., Radhamony,R.,
Thakare,D.R., Bhat,S.R. and Srinivasan.
  TITLE     Cloning of a cryptic promoter driving
expression of reporter gene specifically in roots in
Arabidopsis thaliana
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 446)
  AUTHORS   Sivanandan,C., Prasad,A.M., Radhamony,R.,
Thakare,D.R., Bhat,S.R. and Srinivasan.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-APR-2004) National Research
Centre on Plant Biotechnology, Indian Agricultural
```

Research Institute, Pusa, NewDelhi, Delhi 110 012, India
FEATURES Location/Qualifiers
source 1..446
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="4"
/map="between At4G13930 and At4G13940"
promoter <1..>446
/note="cryptic"
ORIGIN
1 gctatgtatc acccggatgt gaaacattat gggtggagat agactattat aaatttattg
61 aaatatacga ttgttactcg tttaatagca aaagtagtac aatgtatata gtttctatcg
121 agaacaagat ctatttaaaa ttcgaaaagt acatttaaaa ttcataaaca tataaagata
181 gtaacatgtt agatctgcat agtaccacca aaacaagaaa aaagaaacgc acatcgccac
241 ataattgcta tgattctcac tgtcggctgc tttgaaatat tacgattctt ttggtaaatc
301 acacaacata atataattac aataaatata tatatactaa agtataatta atataattaa
361 taccacattg tttaattctg ttttgatctt ttaagatcag tcagatccac cgacgttcct
421 acacgcgcag gtccagatcc aaacag
//

A.11: Sequence of the promoter and UTR of SAHH gene submitted to the Genbank database Accn. No. AY601850

AY601850. Arabidopsis thali...[gi:47174866]
LOCUS AY601850 531 bp DNA linear PLN 19-MAY-2004
DEFINITION Arabidopsis thaliana S-adenosyl homocysteine hydrolase (At4G13940)
gene, upstream regulatory region and 5' UTR.
ACCESSION AY601850
VERSION AY601850.1 GI:47174866
KEYWORDS .
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales;
Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 531)
AUTHORS Sivanandan,C., Prasad,A.M., Radhamony,R., Thakare,D.R., Bhat,S.R. and Srinivasan.
TITLE Cloning and expression analysis of 5' upstream sequence of gene At4G13940 in Arabidopsis
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 531)
AUTHORS Sivanandan,C., Prasad,A.M., Radhamony,R., Thakare,D.R., Bhat,S.R. and Srinivasan.
TITLE Direct Submission
JOURNAL Submitted (17-APR-2004) National Research Centre on Plant Biotechnology, Indian Agricultural Research Institute, Pusa, NewDelhi, Delhi 110 012, India

FEATURES	Location/Qualifiers
source	1..531
	/organism="Arabidopsis thaliana"
	/mol_type="genomic DNA"
	/db_xref="taxon:3702"
	/chromosome="4"
gene	<1..>531
	/locus_tag="At4G13940"
misc feature	<1..468
	/locus_tag="At4G13940"
	/note="upstream regulatory region"
mRNA	469..>531
	/locus_tag="At4G13940"
	/product="S-adenosyl homocysteine hydrolase"
5'UTR	469..>531
	/locus_tag="At4G13940"

ORIGIN

```
        1 tgagtgatac cattagcgcg atacaagcgg gactataaac tgatctagat tgtttttctt
       61 gggaaaaaat gttacaaatt ttaaatatgt agtttgaatt gttaaaccaa gattcaacag
      121 aaatataccg taaataaaca acagttgata atagtcatcg aaaagatatc aactgattct
      181 tcacttgggc tactgtgacg gcccgttagg gttctcaata taagtcaata actacgatct
      241 acgattcact gaaacaaata aaacacagcc acgtgtccac cctcccacat caccgtccga
      301 tctaacccac gacaagctta caacacgggt cataccgctc gtgcagcgtg ttccgtcatc
      361 cacgggatta caacttctac cagatccacc aaaccctcaa aacaatctga accgttcatt
      421 tcattttgac ctcatctata tattctctgt cactcccttt ctcttctcct cgcacacact
```

```
      481 tctctctctc tctctctctg cctcctttcg gattcaaatc
tcagatctag c
//
```