Identifying Insulators in *Arabidopsis thaliana*

Batool Gandorah

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
University of Ottawa
In partial fulfillment of the requirement for the M.Sc. degree in the

Ottawa-Carleton Institute of Biology

Thèse soumise à la
Faculté des études supérieures et postdoctorales
Université d'Ottawa
En vue de l'obtention de la maîtrise ès science

L'Institut de biologie d'Ottawa-Carleton

Abstract:

In transgenic research the precise control of transgene expression is crucial in order to obtain transformed organisms with expected desirable traits. A broad range of transgenic plants use the constitutive cauliflower mosaic virus (CaMV) *35S* promoter to drive expression of selectable marker genes. Due to its strong enhancer function, this promoter can disturb the specificity of nearby eukaryotic promoters. When inserted immediately downstream of the *35S* promoter in transformation vectors, special DNA sequences called insulators can prevent the influence of the CaMV*35S* promoter/enhancer on adjacent tissue-specific promoters for the transgene. Insulators occur naturally in organisms such as yeasts and animals but few insulators have been found in plants. Therefore, the goal of this study is to identify DNA sequences with insulator activity in *Arabidopsis thaliana*.

A random oligonucleotide library was designed as an initial step to obtain potential insulators capable of blocking enhancer-promoter interactions in transgenic plants. Fragments from this library with insulator activity were identified and re-cloned into pB31, in order to confirm their activity. To date, one insulator sequence (CLO I-3) has been identified as likely possessing enhancer-blocking activity. Also, two other oligonucleotide sequences (CLO II-10 and CLO III-78) may possess insulator activity but more sampling is needed to confirm their activity. Further studies are needed to validate the function of plant insulator(s) and characterize their associated proteins.

Résumé

Lorsqu'on effectue une recherche portant sur les transgènes, le contrôle de l'expression génétique est essentiel puisqu'il aide à acquérir des organismes transformés ayant des traits désirables. Plusieurs plantes transgéniques utilisent le promoteur 35S du virus de la mosaïque du chou-fleur, ce qui aide à la sélection des gènes. A cause de la forte activité de son amplificateur, ce promoteur peut interférer avec la spécificité des promoteurs eucaryotiques avoisinants. Des séquences d'ADN appelées isolateurs peuvent empêcher l'influence du promoteur/amplificateur du CaMV35S sur les promoteurs tissu-spécifiques adjacents au transgène, lorsque les isolateurs sont insérés immédiatement en aval du promoteur 35S dans les vecteurs de transformation. Les isolateurs sont naturellement présents dans les organismes telsque les animaux et les levures, mais peuvont été trouvés dans les plantes. Alors, le but de cette étude est d'identifier des séquences d'ADN qui peuvent fonctionner comme isolateurs chez *Arabidopsis thaliana*.

En premier lieu, une bibliothèque d'oligonucléotides synthétisés aléatoirement a été créée à fin d'obtenir des isolateurs potentiels qui seront en mesure de bloquer les interactions amplificateurs-promoteurs dans les plantes transgéniques. Des fragments de cette bibliothèque démontrant une activité d'isolateur ont été identifiés et reclonés dans le vecteur pB31 afin de confirmer leur activité. À ce jour, une séquence d'isolateur (CLO I-3) a été identifiée comme possédant probablement la capacité de bloquer des amplificateurs. De plus, deux autres séquences d'oligonucléotides (CLO II-10 et CLO

III-78) ont été identifiées comme ayant possiblement une activité d'isolateur mais, à fin de confirmer ceci, il sera nécessaire d'obtenir plus d'échantillons. De plus amples d'études seront nécessaires pour valider la fonction d'isolateurs chez les plantes et définir les caractéristiques des protéines qui y sont associées.

Acknowledgments

I would like to thank Dr. Douglas Johnson for the opportunity to work in his lab and all of his guidance and support, and my committee members Dr. Therese Ouellet and Dr. T. Xing. Thanks go as well to Lara Rasooli for her role in the project, working on a different part of a conceptual whole. Special thanks and credit to Dr. Loreta Gudynaite-Savitch, Tatjana Semiz, and Katrina Hiiback; their files and previous work provided the basis for my portion of this study. Also, special thanks goes for Dora Laczko how helped to finish this part of the project, my parents and my husband Mohammad Aburuzizah for all their support and help.

Table of Contents

List of Figures and Tables

List of abbreviations

AGIP promoter	*AGAMOUS* second intron-derived promoter
BLAST	Basic Local Alignment Search Tool
BSA	Bovine serum albumin
°C	Celsius degree
ddH_2O	Distilled de-ionized water
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonuleotide triphosphate, any of the four bases
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
g	Gram
GUS	*β*- glucuronidase
hptII	Hygromycinphosphotransferase type II
Kb	Kilobase
KCI	Potassium chloride
kDa	Kilodaltons
$K_4Fe{\cdot}(CN)_6{\cdot}3(H_2O)$	Potassium ferrocyanidetrihydrate
$K_3Fe{\cdot}(CN)_6$	Potassium ferricyanide
L	Liter
LB	Luriabroth
mg	Milligram
µg	Microgram
$MgCl_2$	Magnesium chloride
$MgSO_4$	Magnesium sulfate
min	Minute
µl	Microliter
mL	Milliliter
µM	Micromolar
mM	Millimolar
NaCI	Sodium chloride
$NaH_2PO_4{\cdot}H_2O$	Sodium dihydrogen phosphate monohydrate
$Na_2HPO_4{\cdot}2(H_2O)$	Sodium phosphate dibasic dihydrate
NaOH	Sodium hydroxide
PBI	Plant Biotechnology Institute (NRC)
PCR	Polymerase chain reaction
rpm	Revolutions per minute
Taq	*Thermus aquaticus*
TBE (10X)	0.89 M Tris, 0.89 M Borate, 20 mM EDTA
T-DNA	Transferred DNA
Tris	Tris (hydroxymethyl) aminomethane
UBI	United Bioinformatica Inc
X-Gluc	5-bromo-4-chloro-3-indolyl-*β*-D-glucuronic acid
35S	Cauliflower mosaic virus 35S promoter/enhancer
35S46	Core cauliflower mosaic virus 35S promoter

Chapter 1: Introduction

A fundamental question in biology is the role played by sequences such as enhancers, silencers and insulators in the regulation of genes. This thesis describes attempts to isolate sequences with insulator activity in the model plant *Arabidopsis thaliana*.

1.1 Chromatin structure and gene regulation

In the eukaryotic nucleus, there is abundant evidence that chromatin not only plays a structural role in the organization of DNA but is also closely involved in the regulation of eukaryotic gene expression (West *et al.* 2002; Litt *et al*. 2001). Several studies have examined the organization of chromatin and found that eukaryotic genomes are divided into functionally autonomous domains in which gene expression is either repressed or facilitated (Wei *et al*. 2005). Active genes or potentially active genes are packaged into a chromatin structure defined as euchromatin or as an open chromatin domain. Euchromatin is a loosely packed form of chromatin that is concentrated with genes accessible to transcriptional enzymes, and is often (but not always) under active transcription. The unfolded structure of the euchromatin domain allows regulatory proteins and RNA polymerase complexes to bind to the DNA sequence, which can then initiate the transcription process at the promoter of an active gene. On the other hand, inactive genes are packaged into heterochromatin, a closed chromatin domain that is transcriptionally inactive, and tightly coiled which makes it less accessible to transcriptional enzymes (West *et al*. 2002).

The basic unit of chromatin is a nucleosome that consists of approximately 147base pairs of DNA wrapped in 1.67 left-handed super helical turns around a histone octamer consisting of 2 copies each of the core histones H2A, H2B, H3, and H4 (Luger *et al.*1997). Modifications to the histone proteins can strongly influence gene expression by what is termed the "histone code". Transcriptionally active euchromatin is associated with hyperacetylation of H3 lysine 9 and 14, while repressive heterochromatin is associated with hypoacetylation of DNA, hyperrmethylation of histone H3 lysine 9 as well as the binding of heterochromatin protein 1 (HP1) (Brasset and Vaury 2005) as can be seen in **Figure 1**.

1.2 Enhancers, silencers and gene regulation

Gene expression is highly regulated. The expression of genes in different tissues at different times of development is controlled by *cis*-acting chromosomal regulatory elements, such as enhancers and silencers. An enhancer is a regulatory sequence located in the euchromatin domain that markedly increases expression of a neighboring gene (Gaszner and Felsenfeld 2006). In contrast, a silencer is a type of regulatory sequence found in the heterochromatin domain that recruits repressor proteins, which are propagated along the DNA in order to repress expression of neighboring genes. For instance, in *Saccharomyces cerevisiae*, silencers in the silenced chromatin recruit Sir proteins that spread along the chromatin producing inactive chromatin domain (Lynch *et al*. 2008). Both enhancers and silencers exert long-distance effects independently of their position and orientation (Brasset and Vaury 2005).

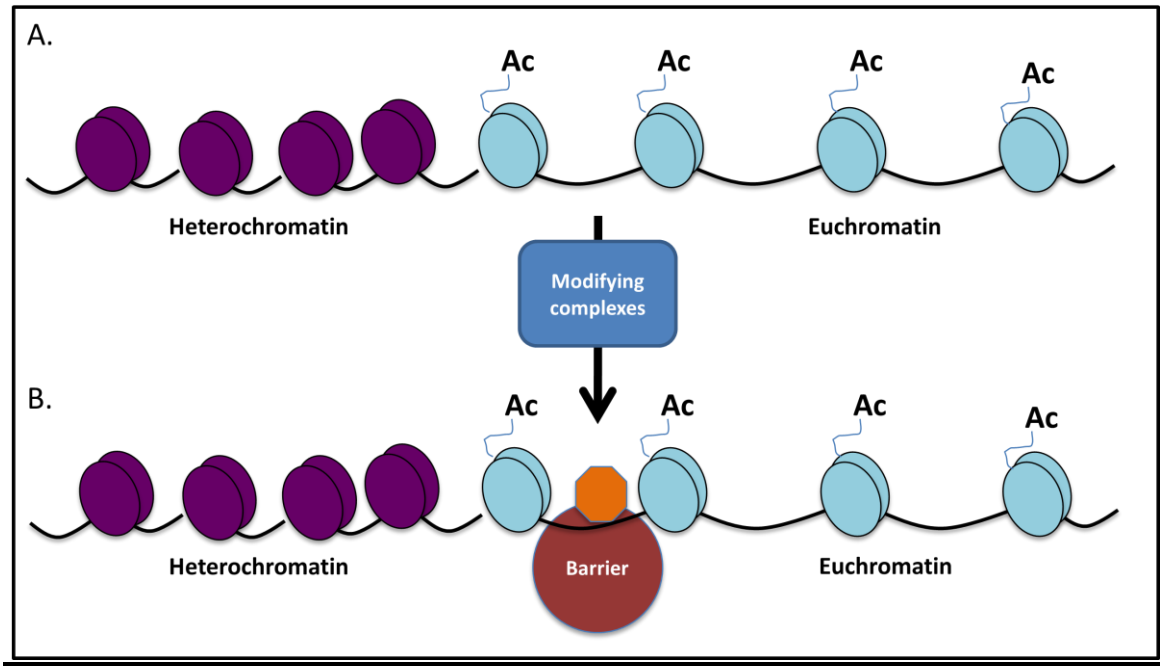


Figure 1

Mechanism underling barrier formation between heterochromatin and euchromatin domains. **(A)** Chromatin divided into two states, heterochromatin and euchromatin and transcriptionally active euchromatin. Modifying complexes can change one state into another. (B) Insulators with barrier activity can change the local balance by assembling or recruiting barrier complexes and prevent unfavourable heterochromatin formation. Ac, acetyl group (modified from Gaszner and Felsenfeld 2006).

Studies have shown that yeast enhancers, which are called upstream activating sequences (UASs) are less flexible in terms of distance and position relative to the regulated promoter compared with their metazoan counterparts, but they can nevertheless work from a distance of up to 1200bp from the promoter (Escher *et al*. 2000). In metazoans, enhancers can influence gene expression independent of their orientation and from remote locations, upstream, downstream or even in introns of the corresponding transcriptional unit (Petrascheck *et al*. 2005). As a result, there is very little correlation between the location of enhancers and that of their target genes within a given genome (Singer *et al.* 2011). For example, the *Drosophila cut* locus is positioned 85-kb upstream of its target promoter (Jack *et al*. 1991). In contrast, enhancers for the rat and human *myosin light chain* 1/3 genes are located downstream of their polyadenylation signals (Donoghue *et al*. 1988) and some enhancers are able to activate transcription from an allelic promoter on a separate chromosome (Morris *et al.* 1998). In plants, studies show that although enhancers are usually found in close proximity to their cognate promoters, there are also some that can function from a distance (Sieburth and Meyerowitz 1997) Furthermore, transgenic *Arabidopsis* bearing the *35S* enhancer situated approximately 3-kb upstream of a target minimal promoter fused to the *GUS* coding sequence displayed constitutive *GUS* expression reminiscent of *35S* promoter-conferred activity indicating that the *35S* enhancer was able to override a 3-kb distance barrier (Singer *et al.* 2010).

Three models have been proposed in order to understand the mechanism of action underlining long-distance enhancer-promoter interactions. While there is evidence supporting each model, there does not appear to be unambiguous proof favoring any one of them (Singer *et al.* 2011). The scanning model (also known as sliding or tracking)

proposes that enhancer-bound activators have a high affinity for the RNA polymerase II complex, or parts of it, and recruit it to DNA. Starting from the enhancer, the recruited complex subsequently scans the DNA until it encounters a competent promoter sequence, where it initiates transcription (Petrascheck *et al*. 2005; Singer *et al*. 2011). Conversely, the looping model suggests that a direct communication through physical apposition between the enhancer-bound and the promoter-bound proteins with concomitant looping out of the intervening DNA causes gene transcription to turn on (Gaszner and Felsenfeld 2006). Finally, the facilitated tracking model, which includes features from both previous models (scanning and looping models), proposed that the enhancer itself and its associated proteins migrate along the intervening chromatin until a stable interaction is established with promoter-bound proteins, resulting in a loop of intervening DNA that grows gradually during the scanning process (Singer *et al*. 2011).

Transcriptionally active genes in the euchromatin domain are surrounded by regions of condensed chromatin (heterochromatin domains) containing silencers. Silencers are initiating elements for silenced chromatin that function by recruiting specific repressor proteins that spread along the DNA producing inactive chromatin domains. These elements could overflow their borders and silence a gene in euchromatin.

1.3 Insulators and gene regulation

The functional flexibility of these regulatory elements (enhancers and silencers) might disrupt normal expression patterns of nearby genes. Since this is rarely observed, mechanisms must exist to ensure that genes are not activated in the wrong tissue or at the wrong time of development by enhancers from neighboring genes or from silencers from

adjacent inactive domains. Over the last few decades, studies have shown that the formation of independent domains of gene function may depend upon specialized DNA sequence elements and their associated binding proteins. Such sequence elements are called chromatin boundaries or insulators. Insulators have been experimentally identified as an additional class of *cis*-acting chromosomal elements capable of preventing inappropriate interactions between adjacent chromatin domains (Gazner and Felsenfeld 2006) and controlling gene expression by regulating interactions between enhancers and promoters. Also, mounting evidence suggests that these sequences play a complex role in the regulation of transcription, perhaps through changes to chromatin structure and nuclear organization (Mongelard and Corces 2001). Experiments that have been done on transgene constructs (mainly in *Drosophila*) have helped to define the general properties of these elements. Studies have shown that even though insulator DNA sequences and their associated proteins are different in different organisms, many exhibit similar properties (Bell *et al*. 2001). For example, *Drosophila gypsy* insulator that binds a group of proteins including a Suppressor of Hairy wing Su(Hw) protein, and the vertebrate hypersensitive site-4 (HS4) insulator that depends on CTCF protein (CCCTC-binding factor) for function; both possess enhancer-blocking activity (Wei *et al.* 2005).

Insulators are typically defined by two functional properties illustrated in **Figure 2a & b**. First, an insulator can act as a barrier preventing the linear spread of nearby condensed chromatin to protect transcriptionally active genes in the euchromatin domain from silencing, possibly by maintaining the histone modification at the boundary between domains or by breaking the code of histone modifications necessary for the propagation of silencing along the chromatin (Brasset and Vaury 2005; Gazner and Felsenfeld 2006).

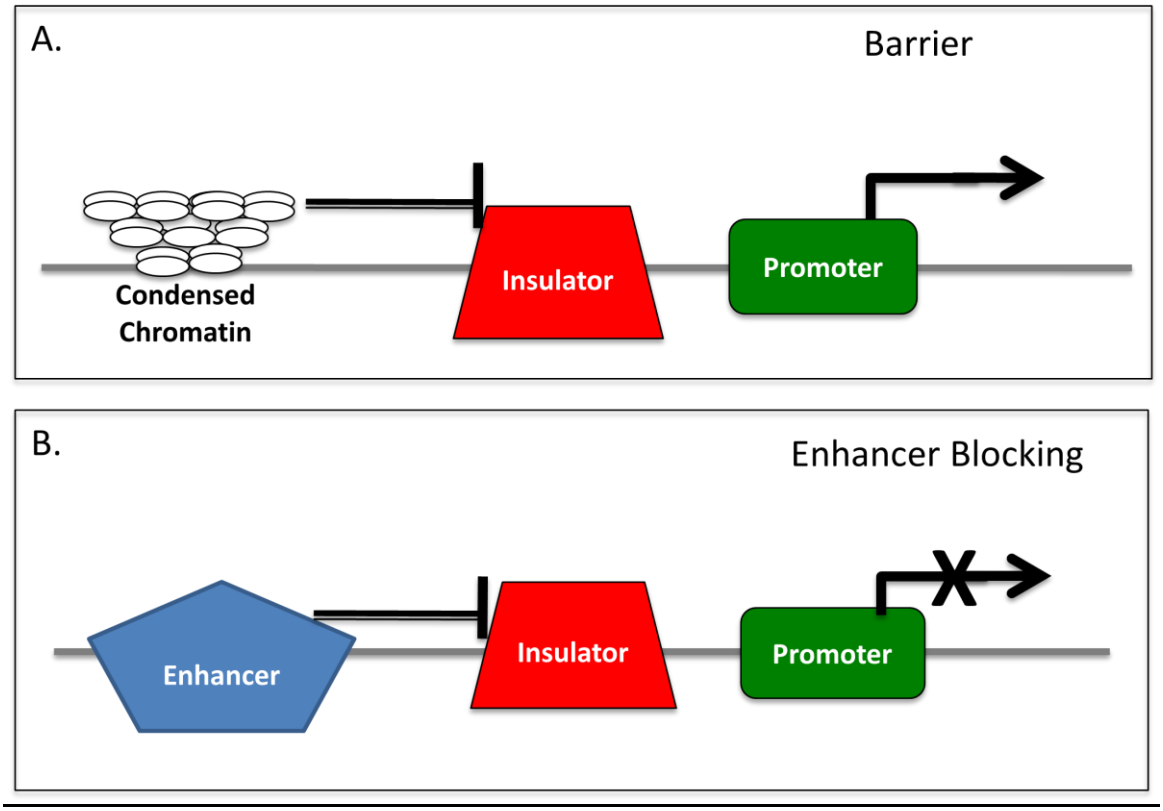


Figure 2

Insulators block enhancer and silencer elements in a position-dependent manner. (A) Barrier elements block the linear spread of silenced chromatin protecting the reporter gene from silencing. (B) Enhancer-blocking elements interfere with enhanced transcription when placed between an enhancer element and its target promoter (modified from Valenzuela *et al*. 2006).

Thus, some insulators do not behave simply as static barriers, as it was first thought, but they can act as a kind of modulator switch which allows them to function as sophisticated regulatory elements, as suggested in (Bell *et al*. 2001). Second, some insulators are able to act as positional enhancer blocking elements **(Figure 2b)**, interfering with the enhancer-promoter interactions only when positioned in between them. True insulators do not affect the function of promoters or enhancers, which indicates their role as neutral blocking elements (Bell *et al*. 2001). Also, some insulators such as compound insulators possess both barrier and enhancer-blocking activities (West *et al.* 2002). Mechanistic differences between the two major types of insulator function at the molecular level remain unclear; however, recent findings suggest that distinct short motifs within each insulator sequence can have different functions and operate independently of each other (Majumder *et al*. 2009), indicating that the sequences and protein co-factors underlying each function may be different.

1.4 Models for insulator function and gene regulation

A number of models for each insulator type have been proposed. For barrier function, two models have been suggested based on the biological mechanism of action of the chicken β-globin insulator in protecting against position-dependent transgene silencing (Mutskov *et al*. 2002; Valenzuela *et al.* 2006). One model posits that the insulator marks the chromatin domains by recruiting chromatin remodeling complexes and histone modifying enzymes such as histone acyltransferase (HAT) at the boundary between domains (Mutskov *et al*. 2002). A second model suggests that the insulators can work as a static barrier by excluding or blocking the access of repressor complexes to the promoter of the flanked transgene; these repressor complexes include histone deacetylases

(HDACs), DNA methyltransferases, methyl-binding domain (MBD) proteins, transcription repressors, and protein such as heterochromatin protein 1 (HP1) that are associated with condensed chromatin (**Figure 2**, Mutskov *et al.* 2002).

Several models such as the transcriptional and the structural models describe enhancer-blocking activity. The transcriptional model supports the idea that insulators have a direct impact on transcription (Geyer 1997; Bell and Felsenfeld 1999). This model can be subdivided into two different models: the promoter decoy model and the facilitator model. The promoter decoy model **(Figure 3a)** postulates that the enhancer advances as an obligatory propagated signal towards the promoter, then an insulator recruits components of the transcription machinery, competes with the promoter for interaction with the enhancer, and minimizes true promoter/enhancer interaction.

Experimental evidence obtained from studies on *Drosophila melanogaster* special chromatin structures (*scs and scs')* suggests that they may contain promoter elements that titrate the enhancer function and keep it from activating transcription (Brasset and Vaury 2005). The facilitator model (**Figure 3b**) suggests that enhancers need "facilitator factors" to be able to interact with the promoter. The insulator interferes with the function of the facilitator factors in preventing the enhancer from effectively communicating with the promoter. In this case, it works as a physical barrier impeding the activation of a gene by its enhancer. This model was supported by experiments performed on *Drosophila melanogaster* transcription factor GAGA. This transcription factor stimulates transcription by linking an enhancer to its cognate promoter thus providing a protein bridge that mediates enhancer-promoter communication. However, the presence of insulators could impede the function of GAGA, and restrict the recruitment of this factor to the promoter (Mahmoudi *et al.* 2002).

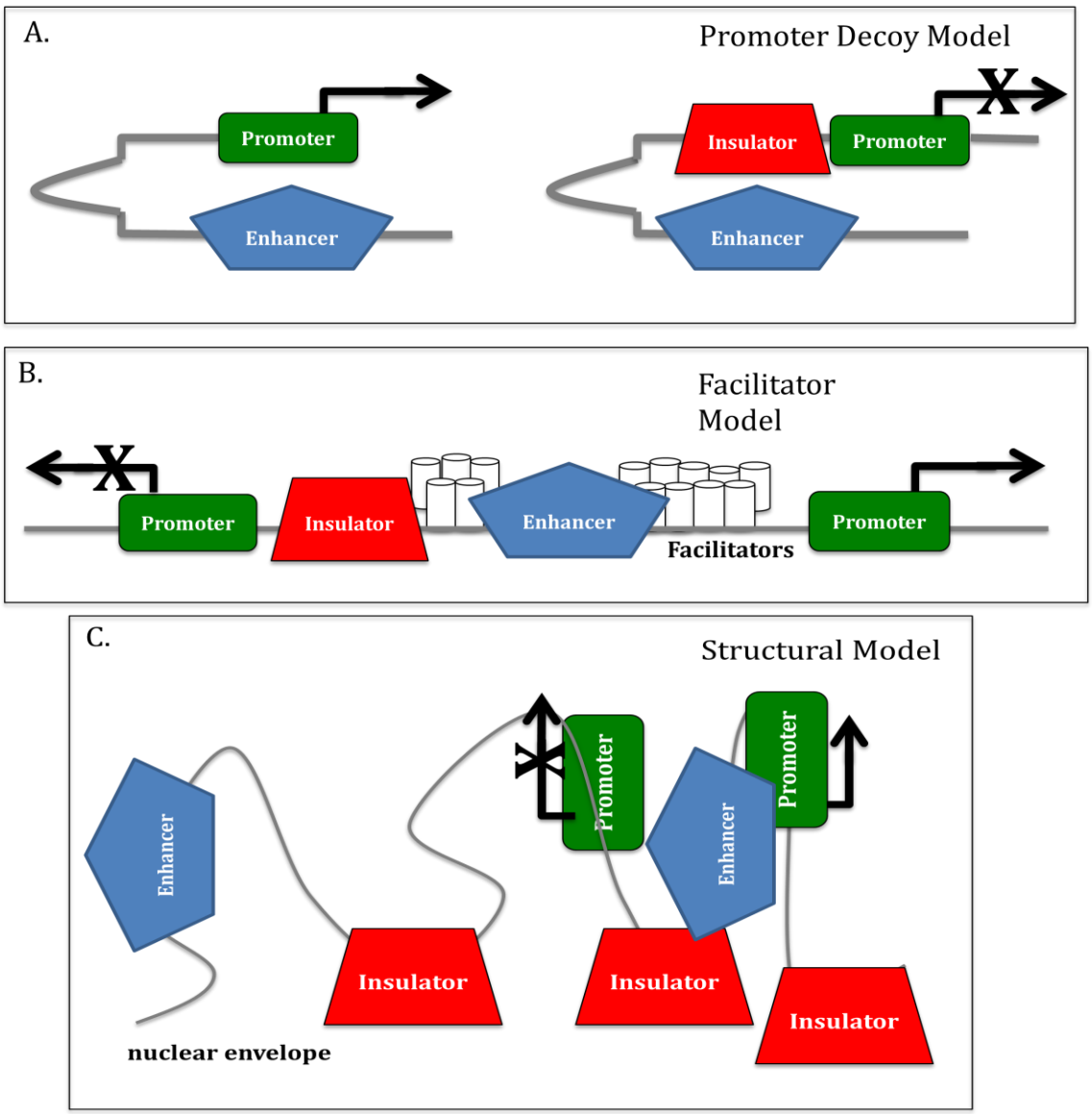


Figure 3

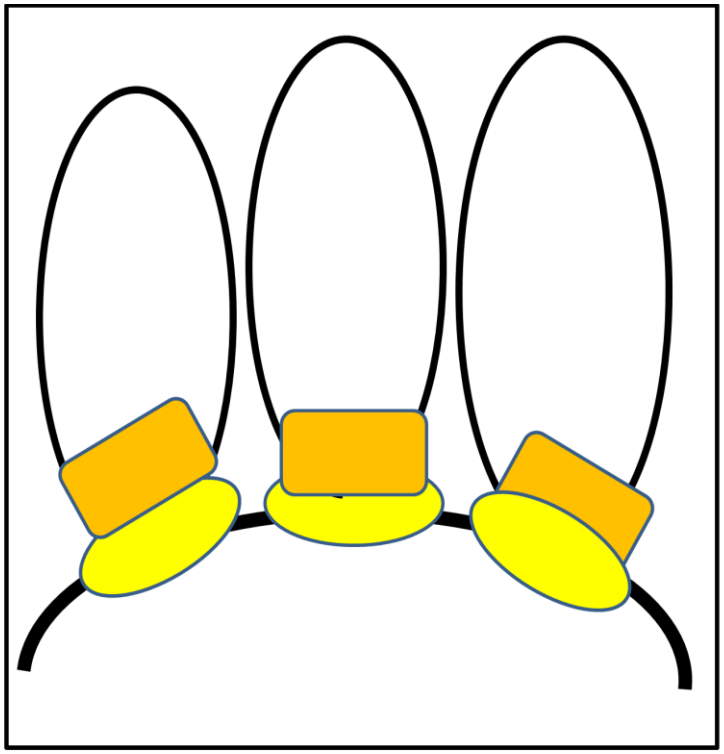
Enhancer-blocking models. (A) Decoy model: insulator protein complex directly interacts with the enhancer-bound proteins intercepting the enhancer signal into a non-productive interaction.(B) Facilitator model, the insulator interferes with the function of facilitator proteins to prevent the enhancer from communicates with the promoter. (C) Structural model, insulator organizes the chromatin fiber into loops. Each loop represents an independent structural and functional domain, and the interaction between the enhancer and the promoter only occurs if they are located in the same loop (modified from Valenzuela *et al*. 2006).

In fact, both of these transcriptional models seem too simplistic and limited because neither one can explain why boundary elements have to be placed between the enhancer and the promoter to function as enhancer-blocking elements. Furthermore, it also fails to explain how an insulator blocking an enhancer on one side can activate a promoter on its other side (Brasset and Vaury 2005; Gaszner and Felsenfeld 2006). Studies have shown that another model involving proteins that aid in bringing the enhancer and the promoter close to each other might be considered as an alternative model such as the loop or structural model. This model proposed that the interaction between a pairing of insulator elements or between an insulator and a fixed structure (nuclear lamina) result in organization of chromatin into loops with distinct functional domains (Bushey *et al*. 2008), which isolate the signal generated in a domain.

As a result an enhancer can interact with its promoter only if they are located in the same loop (**Figure 3c**) (Valenzuela *et al*. 2006). An example of this structural model, is the interaction between the *Drosophila* Su(Hw) insulator and Chip protein that facilitates the formation of clusters that bring enhancers and promoters closer together, causing the Su(Hw) insulator to interfere with the role of this protein (Gaszner and Felsenfeld 2006).Genetic evidence has shown that Su(Hw) becomes a stronger insulator when enhancer-promoter communication is weakened by mutations in Chip. The mutations seem to cause the formation of Chip-Su(Hw) complexes to break the chain of interaction between Chip and homeodomain proteins (Brasset and Vaury 2005). In addition, the observation that *gypsy* insulator complexes cluster in the nucleus supports this model (Capelson and Corces 2005).

All proposed models for the mechanism of insulator activity require both *cis*- and *trans*-acting elements working together; different *cis* insulator DNA sequences bind different *trans* facilitating proteins such as CTCF in vertebrates. These proteins recruit other special proteins that facilitate subsequent attachment of the insulator to the nuclear lamella to create chromatin loops called "insulator bodies" **(Figure 4)**. These bodies function by forming higher-order chromatin structures and insulator effect, with each component protein required for activity (Ong and Corces 2009). Thus, it seems that the formation of the chromatin loops depends on these specific *trans*-acting proteins because loss of function or knockout mutations in the proteins can eliminate insulator function (Golovnin *et al*. 2008).

Studies have shown that barrier and enhancer-blocking activities are separate since each one depends on specific DNA sequences for function. Functional mapping of the *Drosophila* SF1 insulator which is a 2.4-kb boundary element located in the intergenic region between the non-Hox gene *fushi tarazu* (*ftz)* and the homeotic gene Sex comb reduced (*Scr*) in the *Drosophila* Antennapedia homeotic complex (ANT-C) (Belozerov *et al*. 2003; Majumder *et al*. 2009) has revealed that different DNA elements mediate the enhancer-blocking and barrier activities of insulators. For example, the *Drosophila* SF1 insulator sub-fragment SF1b that possesses a strong enhancer-blocking activity is not able to protect a transgene from chromosomal position effect, while the sub-fragment SF1c, which has no enhancer-blocking activity, has strong barrier activity (Cai *et al*. 2003). Also, each function requires a distinct sequence: SF1b enhancer-blocking activity depends on the GAGA sites whereas the barrier function of FS1c appears to be GAGA-independent (Cai *et al*. 2003).



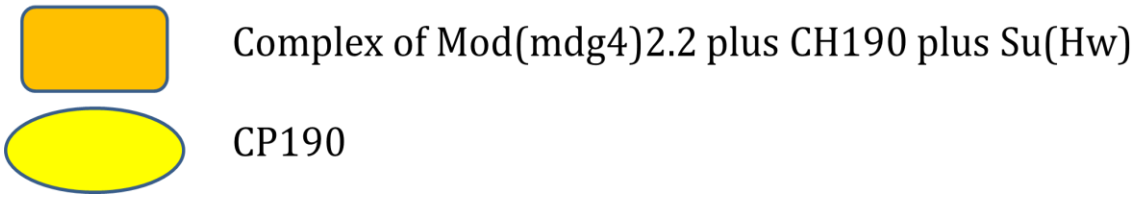


Figure 4

Insulator body formation in *D. melanogaster* by the *gypsy* insulator. Insulator bodies are formed by the association of the Suppressor of Hairy wing Su(Hw) protein with other specific DNA elements, and other proteins ,through its zinc fingers, including: the centrosomal protein 190 kD ,CP190 and modifier of mdg4, Mod(mdg4)2.2; the latter may associate with topoisomerase-I-interacting protein,dTopors, bringing the insulator body to the nuclear lamina. (Modified from Gaszner and Felsenfeld, 2006; Ong and Corces, 2009).

In addition, both barrier and enhancer-blocking insulators have the ability to reduce chromosomal position effect (CPE) when they flank a transgene in a recipient genome. CPE refers to variability noted in transgene expression levels when the gene is moved away from its native context to a new chromosomal environment. In transgenic research the precise control of transgene expression is crucial in order to obtain a transformed organism with expected desirable traits; therefore, position effect is a big concern because it can strongly influence the transcription of foreign genes in transgenic organisms resulting in low frequencies and levels of gene expression and, in some cases, in aberrant patterns of expression.

Data from all characterized insulators possessing barrier and/or enhancer-blocking activities thus far seem to support the proposed looping model with defined chromatin domains bringing some elements into proximity, allowing activation of some genes, or isolating others, causing repression (Mongelard *et al*. 2001; Valenzuela *et al*. 2006). The nature of this looping model can support both insulator activities since it can provide a physical separation required for both types of insulator function. In addition, the protein bodies that mediate this higher order structure intuitively seem to act as crucial elements maintaining the functional domains, confirmed by the loss of insulator activity in the protein's absence (Majumder and Cai 2003; Valenzuela and Kamakaka 2006; Gazner and Felsenfeld 2006).

Data obtained from various studies that have been done in different organisms is consistent with the looping structure model hat is associated with different mediating proteins. For instance, in *Drosophila*, the *gypsy* insulator directly binds Su(Hw) protein. In turn this protein utilizes 12 zinc fingers to bind two additional proteins, centrosomal

protein 190 (CP190) and the protein modifier of mdg4, both necessary for loop domain formation to form insulator bodies, as shown in **Figure 4** (Gaszner and Felsenfeld 2006).

In vertebrates, CTCF helps to displace insulators to the periphery of the nucleoli generating similar loops described for the *gypsy* insulator (Brasset and Vaury 2005). In *A. thaliana*, a similar mechanistic effect can be observed in the regulatory silencing of*KNOTTED1*-like homeobox *KNOX* genes during organogenesis when two proteins, ASYMMETRIC LEAVES1 AS1 and AS2, interact with each other and form a repressor complex that binds directly to two DNA motifs that flank the enhancer element of the *KNOX* genes *BREVIPEDICELLUS* (*BP*). Recruitment of the histone chaperone histone gene repressor A(HIRA) is necessary for this process, and it may act by facilitating looping of the enhancer element**(Figure 5)** (Ong and Corces 2009). This does not represent a true insulator but it suggests that plants can use the looping mechanism to regulate their gene expression. Thus, the distribution of this mechanism among different organisms therefore supports its importance and conservation and further suggests a role for insulators in genetic regulation (Ong and Corces 2009).

Investigating the enhancer-blocking activity of insulators on transgenic constructs has revealed that pairing between two homologous (Cai and Shen 2001) or heterologous Su(Hw) insulators, or more specifically when pairing occurs between the binding sites for the GAGA factor and the *gypsy*, insulator activity is abolished (Melnikova *et al*. 2004). This bypass probably relies on the bound protein complexes that form a mini-loop to bring the enhancer and promoter into close proximity (Wei *et al.* 2005). In contrast, testing the enhancer-blocking activity of the other eleven selected pairs of *Drosophila* insulators including *scs*, Fab-7, SF1 and ftz-MAR found that unlike Su(Hw), whose activity is

abolished by homologous pairing, the heterologous and even homologous pairing of other *Drosophila* insulators does not abolish their insulator activity (Majumder and Cai 2003). Instead, their enhancer-blocking activity was improved, indicating that diverse mechanisms may underlie insulator function, and selective interaction between different classes of insulator elements may occur (Majumder and Cai 2003).Moreover, studies done by Cai and Shen (2001) using a series of transgenic assays to assess the function of different combinations of enhancers and *gypsy* insulators in *Drosophila* embryos proved that an enhancer was blocked more effectively from a promoter by two flanking Su(Hw) insulators than by a single intervening one.

This demonstrates that the inhibition of transcription occurred due to a change in the organization of the chromatin structure by looping out the chromatin forming distinct functional domains, rather than interfering with enhancer–promoter communications (Mongelard and Corces 2001). Furthermore, insulator effectiveness in transgenic constructs may also vary according to interaction with different, promoter/enhancer systems used (Gudynaite-Savitch *et al*. 2009).

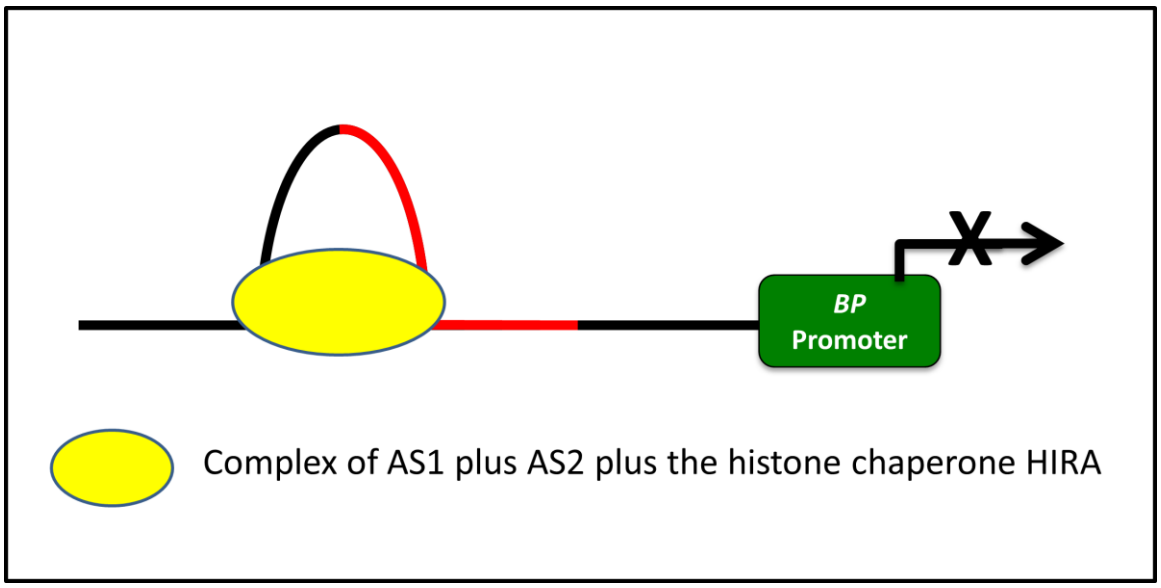


Figure 5

Loop formation in *Arabidopsis* mediated by protein cofactors, which divide chromatin into distinct domains. In *Arabidopsis*, binding of the AS1-AS2 complex flanking the *BP* gene's enhancer is required for silencing, and the recruitment of the histone chaperone HIRA is also necessary for this process since it probably acts by facilitating looping of the enhancer element (modified from Ong and Corces 2009).

Organism	Insulator name	Locus	NCBI accession	Binding Protein	References
Drosophila melanogaster	*gypsy retrotransposon*	no specific locus	Not found	Su(Hw) (Suppressor of Hairy Wing)	Kuhn *et a* l. 2003; Bell *et al* .1999; Valenzuela *et al* . 2006; Mongelard and Corces 2001; Baxley *et a* l. 2011
Drosophila melanogaster	*scs* (*special chromatin structures*)	87A7(*hsp70*) locus *(heat-shock* genes *(hsp70)* on the 87A7 locus)	K01292	Zw5 (zest-white-5)	Kuhn *et al* . 2003; Bell *et al* . 1999; Valenzuela *et al* . 2006
Drosophila melanogaster	*scs'* (*special chromatin structures*)	*87A7(hsp70) locus (heat-shock genes (hsp70)* on the 87A7 locus)	K01292	BEAF (boundary element-associated factor)	Kuhn *et a* l. 2003; Bell *et a* l. 1999; Valenzuela *et a* l. 2006
Drosophila melanogaster	Fab7 (Front abdominal-7)	*Abd-B (Abdominal-B)*	U31961	Not defined	Bell *et a* l. 2001; Valenzuela *et al* . 2006; Kyrchanova *et al* . 2011
Drosophila melanogaster	Fab8 (Front abdominal-8)	Abd-B (*Abdominal-B*)	U31961	CTCF (CCCTC-binding factor or 11-zinc finger protein)	Holohan *et* al. 2007, Bell *et a* l. 2001; Valenzuela *et* al. 2006; Mongelard and Corces 2001; Kyrchanova *et al* . 2011
Drosophila melanogaster	FS1	*ftz-Scr* (region between fushi tarazu (ftz) and Sex comb reduced (Scr)	Not found	CTCF (CCCTC-binding factor or 11-zinc finger protein)	Cai *et al.* 1999; Belozerov *et al* . 2003; Majumder *et al* . 2009
Drosophila melanogaster	Mcp (Miscadastral)	Abd-B (*Abdominal-B*)	U31961	CTCF (CCCTC-binding factor or 11-zinc finger protein)	Holohan *et* al. 2007; Kyrchanova *et al* .2011
Ashbya gossypii	UASrpg (upstream activation sequence of ribosome protein genes)	None	Not found	Rap1p-binding sites (repressor activator protein 1-bining sites)	BI and Broach 1999; Gudynaite-Savitch *et al* . 2009; Idrissi *et al* . 1998; Eriksson *et al* . 2000
Saccharomyces cerevisiae	HML	*HML* locus	Not found	Rap1p-binding (repressor activator protein 1-bining sites)	Fourel *et al* . 1999
Saccharomyces cerevisiae	HMR	*HMR* locus	Not found	Rap1p (repressor activator protein 1-binding sites); Abf1p, Sir1p (silent information regulator); and ORC (origin recognition complex)	Donze *et* al. 1999; Ruben *et al* . 2011; Bose *et al* . 2004
Mus musculus	ICR (imprinted control region)	*lg f2-H19* locus (lg f2 (insulin-like growth factor 2-H19 locus)	Not found	CTCF (CCCTC-binding factor or 11-zinc finger protein)	Bell *et a* l.1999; Valenzuela *et a* l. 2006; Mongelard and Corces 2001
Gallus gallus	HS4 (chicken hypersensitive site-4)	*B-globin* locus	Not found	CTCF (CCCTC-binding factor or 11-zinc finger protein)	Bell *et a* l. 2001; Valenzuela *et a* l. 2006; Yusufzal *et* al. 2004
Homo sapiens	BEAD-1 (blocking element alpha/delta)	human TCRα/ δ (T cell receptor α/δ locus)	NG_001333	CTCF (CCCTC-binding factor or 11-zinc finger protein)	Bell *et a* l. 2001, Valenzuela *et al* .2006; Ramezani *et a* l. 2008; Zhong and Krangel 1997; Sleckman *et al* . 2001; Gudynaite-Savitch *et al* . 2009
Homo sapiens	BEAD-1C (blocking element alpha/delta)	human TCRα/ δ (T cell receptor α/δ locus)	Not found	CTCF (CCCTC-binding factor or 11-zinc finger protein)	Bell *et* al. 2001; Valenzuela *et a* l. 2006; Gudynaite-Savitch *et al* . 2009
Homo sapiens	5'HS4 (5 end hypersensitive site-4)	*5' B-globin* locus	AF064190	CTCF (CCCTC-binding factor or 11-zinc finger protein)	Bell *et al* . 2001; Yoa *et al* . 2003
Bacteriophage Lambda	λ (*EXOB*) bacteriophage lambda fragment	None	Not found	Not defined	Singer *et al.* 2009
Petunia hybrida	*TBS* (transformation booster sequence)	None	EU864306	Not defined	Hily *et al.* 2009; Galliano *et al* . 1995

Table 1: Insulators that have been identified and characterized in various organisms, and their associated binding proteins.

Increasing evidence suggests that insulators may play an important role in controlling gene expression pattern during development by preventing inappropriate enhancer-promoter interactions between adjacent chromatin domains. This occurs by looping out chromatin, generating distinct functional domains and maintaining this physical structure overall which then helps to define different cell types.

1.5 Insulator sequences and their binding partners

Insulators and their associated binding proteins have been identified in many organisms as summarized in **Table 1**. Insulator elements were first discovered in *Drosophila* and subsequently found in various organisms including yeast, sea urchin, chicken, mouse, and human (Wei *et al*. 2005). The *Drosophila gypsy* insulator is one of the most widely studied since it was the first identified insulator within the genome of the *gypsy* retrotransposon. It contains a 340bp sequence with 12 binding sites for DNA binding protein Su(Hw). It confers its activity through a protein complex that consists of three components, Su(Hw), Mod(mdg4)2.2, and CP190. Also, this protein complex interacts with topoisomerase I-interacting RS protein (dTopors) that appears to be involved in the establishment of chromatin organization through its ability to mediate the association of insulator complexes with a fixed nuclear substrate (Capelson and Corces 2005). Enhancer-blocking activity in insulators was first discovered through a mutation in the *Drosophila yellow* gene, which is controlled by multiple tissue-specific enhancers located both upstream and downstream of its promoter. The *yellow* gene is required for dark pigmentation of the *Drosophila* larval and adult cuticle. An experiment showed that insertion of the *gypsy* insulator upstream of the promoter of the *yellow* gene prevents the downstream enhancers

from activating the promoter, resulting in *yellow* instead of dark pigmentation of body and wing cuticle (Bell *et al*. 2001).

Additional insulator systems in *Drosophila* have been studied. DNA in the Bithorax complex near the *abd-B* gene also illustrated insulator enhancer-blocking activity. The *abd-B* gene is required for identification of posterior abdominal segments, and it is controlled by a series of parasegment-specific enhancers. Three elements that function as enhancer-blocking boundaries have been identified in the *abd-B* locus: Miscadastral (MCP) between *iab-4* and *iab-5*, *Frontabdominal-7* (*Fab-7*), between *iab-6* and *iab-7*, and *Fab-8*, between *iab-7* and *iab-8* **(Table 1)**. Mutations in these elements negatively affect their insulator activity resulting in the transformation of one parasegment into another (Bell *et al*. 2001; Valenzuela *et al.* 2006). *Scs* and *scs*' **(Table 1)** elements in *Drosophila* are insulators first identified as sequences possessing barrier activity. They are located ~15-kb apart and mark the ends of the heat-shock genes (*hsp70)* on the 87A7 locus. The two elements contain binding sites for proteins zest-white-5(Zw-5) and boundary element-associated factor (BEAF), respectively (Valenzuela *et al*. 2006). Their ability to protect against position effects was subsequently proven (Wei *et al*. 2005; Bell *et al*. 2001). An experiment where the eye-color gene in fruit flies is flanked by *scs* and *scs'* have shown that these elements caused a reduction in the variability in eye color that appears as a result of position effects (Bell *et al.* 1999).

In vertebrates, the first well characterized insulator was chicken hypersensitive site-4 (cHS4) **(Table 1)**, which is located near the 5' end of the chicken *β-globin* locus. Similar to the *Drosophila* Su(Hw) and *scs* insulators, the cHS4 element functions as an enhancer blocker element and as a barrier to protect transgenes from position effects in *Drosophila*

and in early–erythroid chicken cell lines. However, its barrier and enhancer-blocking functions are separate because each function is regulated by different proteins (Wei *et al*. 2005). Studies in mammals have shown that a ubiquitous mammalian bHLH-ZIP transcription factor (USF) is commonly involved in the insulator's barrier activity, while CTCF is highly associated with the insulator's enhancer blocking activity (Arumugam *et al.* 2009). This insulator resembles a second insulator located at the 3' end of this locus with which it shares the same properties of establishment and/or maintenance of the chromatin domain boundary, and block enhancer-promoter activity (Bell *et al.* 2001).

Additionally, the 1.6-kb BEAD (blocking element alpha/delta, **Table 1**) element derived from the human TCR α/δ (*T cell receptor* α/δ) locus (Zhong and Krangel 1997) functions as an enhancer-blocking element by separating the differentially regulated TCRα and TCRδ genes from each other's enhancers (Bell *et al*. 1999). The imprinted control region (ICR, **Table 1)** found in the mouse *lg f2* (insulin-like growth factor 2)-*H19* locus with four binding sites for the CTCF protein, also works as enhancer-blocking element by repressing the expression of the *lgf2* gene in the maternally transmitted allele (Valenzuela *et al*. 2006). Experimental results showed that methylation of the ICR on the paternal allele could somehow cancel its insulator activity, allowing the downstream enhancer to activate the expression of *lg f2* (Bell *et al.* 2001).

Almost all vertebrate enhancer-blocker insulators characterized need binding of the regulatory protein CTCF for their activity, and it is the only identified trans-acting factor that confers enhancer-blocking insulator activity (Bao *et al.* 2008). CTCF contains 11-zinc finger DNA-binding domains and binds to a wide range of *cis*-elements by utilizing

various combinations of individual zinc fingers (Wei *et al*. 2005). It is an 82 kDa ubiquitously expressed nuclear protein that has been found to be commonly distributed among vertebrate genomes (Filippova *et al.* 1996; Klenova *et al*. 1993).CTCF is characterized by an unusually extensive DNase I footprint (51-bp) when bound to its site on DNA, consistent with an involvement of several fingers in typical binding sites (Bell *et al*. 1999). The fact that the binding sites of this protein have been found to function in various species (Moon *et al*. 2005)indicates that it is an essential (Fedoriw *et al*. 2004) and highly conserved element of at least a proportion of insulator systems (Singer *et al.*2011). Studies have revealed that CTCF has two transcription repressor domains that can act as a transcriptional repressor such as those involved in regulation of gene imprinting in the mouse *lg f2/H19* locus (Bell *et al.* 2001; Kim *et al.* 2007). CTCF may also function as a transcriptional activator in different sequence contexts (Bell *et al* .1999). It seems that CTCF directs the enhancer-blocking activity in different species by using the same model that is proposed for Su(Hw) in *Drosophila*: CTCF molecules can interact with each other to form clusters and therefore create closed loop domains. Also, it has been proposed that CTCF can bind the chromatin at the nucleolar surface through interactions with the nucleolar proteins such as nucleophosmin/B23, a protein that is present throughout the nucleus but is largely concentrated on the nucleolar surface (Yusufzai and Felsenfeld 2004),generating open loop domains (Gaszner and Felsenfeld 2011). Identifying the CTCF site after whittling down the chicken β-globin insulator to a minimal active element really supports the idea that CTCF is responsible for directing the enhancer-blocking activity in vertebrates. Moreover, mutations in the zinc fingers domains of CTCF abolish its ability to confer the enhancer-blocking activity of insulators (Graham *et al*. 2004; West *et al.* 2003).

Recent studies imply that even in those cases where only CTCF sites are present, the activity of CTCF might require the participation of other proteins, like the directional enhancer blocking activity of the Su(Hw) protein, which involves interaction with other proteins such as Mod (mdg4) protein (Bell *et al.* 1999). Because CTCF is an 11-zinc finger protein, therefore, different binding sites in DNA may engage different subsets of fingers, which give each interaction a special feature. These characteristics of the binding site would have a large influence on the conformation of the protein, the nature of its interactions with cofactors, and its ultimate biological effect (Bell *et al*. 1999; Gaszner and Felsenfeld 2011). In vertebrates, experiments have proved that the CTCF site is unnecessary to protect against position effect since it fails to protect a transgene that is located on the inactive mouse X chromosome from being silenced. CTCF also fails to function as a barrier in a chicken cell-based transgene assay (Recillas-Targa *et al.* 2002; Gaszner and Felsenfeld 2011).

1.6 Insulators in plant systems

In plants, few sequences have been identified with consistent insulator activity in an *Arabidopsis thaliana* model system. *A. thaliana* is a small flowering plant native to Europe, Asia, and northwestern Africa commonly used in plant genetics due to its small genome with about 157 mega-base pairs and five chromosomes, a fully sequenced genome, and ease of transformation using *Agrobacterium tumefaciens* T-DNA vectors by the floral dip method.

NI29, a 16bp consensus sequence from *A. thaliana* derived from bioinformatics analysis as described in patent literature (Gan and Xie 2002), was the first insulator

proposed to eliminate inappropriate enhancer-promoter interactions in transgenic plants (Gan and Xie 2002). However, NI29 failed to block the CaMV (cauliflower mosaic virus) *35S* enhancer-mediated activation of adjacent promoters, but instead increased misexpression of transgene (Gudynaite-Savitch *et al.* 2009). Recently, a 2 kb matrix attachment region (MAR) called a transformation booster sequence (*TBS*, **Table 1**) from the plant *Petunia hybrida* was found to function as an enhancer-blocking insulator when inserted between the CaMV*35S* promoter/enhancer and an *AGIP* promoter (*AGAMOUS* second intron-derived promoter) driving the expression of the *GUS* (*β-glucuronidase*) gene in *A. thaliana* (Hily *et al.* 2009).

There is increasing evidence for the involvement of matrix attachment regions (MARs), which are characterized by their ability to bind a network of non-histone proteins located in the nuclear matrix, in the regulation of gene expression since they are another type of DNA sequence that has the ability to organize chromatin into loop domains (Hily *et al.* 2009). Furthermore, experiments show that expression levels of some genes are altered depending on their position relative to the matrix (Anthony and Blaxter 2007). MARs have also been associated with enhanced and more stable transcription, notably when they are positioned on either side of a transgene (Allen *et al.* 2000). It seems that the enhancer-blocking activity of the *TBS* MAR does not only rely on its ability to interact with the nuclear matrix but probably, on an unidentified functional motif within the 2-kb DNA fragment (Hily *et al.* 2009). Thus, these findings suggest a functional connection between insulators, the nuclear matrix and nuclear organization (Brasset and Vaury 2005).

Moreover, a 1.3 kb matrix attachment region from the *β–phaseolin* gene of *Phaseolus vulgaris*, as well as 1.2 kb genomic and 0.9 kb coding sequence fragments from the same

gene have been found to block the activity of the use CaMV*35S* promoter/enhancer on nearby promoters (van der Geest and Hall 1997; Singer *et al*. 2010).

A study that was carried out to test the ability of various bacteriophage λ fragments to block enhancers of target promoters in *A. thaliana* showed that a 1-kb bacteriophage lambda fragment (*EXOB*, **Table 1**) functions as an enhancer-blocking insulator in *A. thaliana* (Singer *et al.* 2010). This experiment was performed by inserting three different bacteriophage λ fragments with different lengths; 1, 2, and 4-kb respectively, between the *35S* enhancer and an *AGIP* promoter fused to GUS. Subsequently, the activity of the *GUS* gene was detected in transgenic leaves. It appeared that neither the 2-kb nor the 4-kb λ fragments were able to block the influence of *35S* on the promoter of the reporter gene whereas the 1-kb λ fragment resulted in diminished expression of the *GUS* gene in transgenic leaves, indicating that there is no correlation between the fragment's length (up to 4 kb) and insulator function (Singer *et al*. 2010). It seems that both a *TBS* and a 1-kb λ fragment function as true insulators because they do not possess silencing activity in the system they were tested (Singer *et al.* 2010).

Due to the lack of characterized plant insulators, elements characterized in other species or described in the literature with enhancer-blocking activity have been tested in plants, some with significant activity (Gudynaite-Savitch *et al.* 2009). For example, BEAD-1 and BEAD-1C insulators **(Table 1)** from the human *T cell receptor* α/δ locus, as well as UASrpg (upstream activation sequence of ribosome protein genes) from the filamentous fungus *Ashbya gossypii* **(Table 1)** (Bi and Broach 1999) were able to reduce non-specific enhancer-promoter interactions when they were positioned between the *35S* enhancer/promoter and transgene promoters in *A. thaliana* (Gudynaite-Savitch *et al.* 2009;

Singer *et al*. 2010). The fact that heterologous DNA provides at least partial blocking of *35S* enhancer-mediated activation in transgenic plants implies that the insulator machinery in this wide range of eukaryotic organisms might be evolutionary conserved since insulator function requires the interaction between *cis*-DNA elements and specific *trans*-proteins (Wallace and Felsenfeld 2007;Gudynaite-Savitch*et al.* 2009; Singer *et al*. 2010). This also suggests that insulators do function in gene regulation in plants, similar to other eukaryotes. In addition, the similarity between animal and plant cells thus hints that the spatial and temporal control of the gene expression in plant cells might also engage DNA sequences with enhancer blocking activity. Furthermore, it has been proven that *A. thaliana* uses chromatin loops as a regulatory feature to regulate the silencing of *KNOX* genes (Ong and Corces 2009). Therefore, it was proposed that plants must possess *cis*-acting sequences that interact with *trans*-acting factors with insulator activity. Finding a *TBS* MAR, which provides one identified example endogenous to a plant species with enhancer-blocking activity, supported this hypothesis as does the recent identification of a gypsy-like element from *A. thaliana* that exhibits insulating activity in a transgenic system (Singer and Cox 2012). Both DNA and protein BLAST searches in *A. thaliana* were unable to find a sequence for CTCF suggesting that different proteins and sequences that carry out the same molecular interactions must substitute for CTCF found in animals. Therefore, further research is needed to identify and characterize more DNA insulator sequences in plants, to eventually characterize their associated proteins and modes of function.

Identifying such sequences will help to elucidate the true mechanisms behind both enhancer-promoter, and enhancer-insulator interactions. Also, it would help to

understandthe mechanism by which genes are regulated overall in plants. In addition, since both insulator functions are able to produce a reduction in position effect by flanking a transgene, finding plant sequences with insulator activitywould provide a powerful tool in plant biotechnology. This would allow for strict control of tissue-specific transgene expression allowing engineering for more stable and predictable transgenic organisms.

In plants, genetic engineering research has played a significant role in genetic improvement of agronomically important traits in crops. Most transgenic research that has been conducted has involved the improvement of a single trait; however, plants in field conditions must withstand different kinds of biotic and abiotic challenges (Singer *et al.* 2011). This issue can be solved through using transformation vectors that harbor multiple transcriptional gene units to enhance several traits in the crop at the same time (Gudynaite-Savitch *et al.* 2009). The presence of multiple enhancer/promoter elements within a single vector often leads to inappropriate interactions between the enhancers and nearby promoters of other genes, as is the case in transgenic plants with vectors using constitutive promoters to drive the expression of selectable marker genes such as *hygromycin phosphotransferase II (hptII)*. Constitutive promoters such as CaMV*35S* enhancer/promoter can affect the correct tissue-specific expression of transgenes due to its strong enhancer function resulting in disturbances of the specificity and consistency of its transcriptional activity in transgenic plants.

Several strategies that might reduce or eliminate inappropriate enhancer-promoter interactions were tested including transgene configuration, spacer DNA, and the replacement of the CaMV*35S* promoter/enhancer (Gudynaite-Savitch *et al.* 2009). Certainly, a number of studies have proven that increasing the distance between enhancer

and promoter in transgenic constructs reduces interference. Two factors must be considered for determining the length of the DNA spacer required, the strength of the enhancer and the sensitivity of the target promoter (Singer *et al.* 2011). For instance, a 2.7-kb fragment was inserted between the *35S* enhancer and *napin* (the *Brassica napus napA* seed-specific promoter), and between the same enhancer and *prx* (soybean seed coat hour-glass cell peroxidase promoter), respectively; enhancer activation was only observed in the case of *prx* (Gudynaite-Savitch *et al.* 2009; Singer *et al*. 2011). Since, inserting large spacer fragments into transformation vectors could cause technical problems for cloning and vector stability (Singer *et al*. 2011) this strategy is not usually efficient for use in transgenic research.

Recent efforts have focused on the use of tissue-specific promoters that restrict the transgene expression to targeted tissues only, in order to avoid the crosstalk between CaMV*35S* enhancer and nearby promoters (Gudynaite-Savitch *et al*. 2009). For example, during fruit ripening, E-8-fruit-ripening-specifc polygalacturonase and fruit-specific phytoenedesaturase promoters have been used to generate tomato plants with delayed ripening or modified carotenoid content (Rosati *et al*. 2000; Fraser *et al.* 2002; Gudynaite-Savitch *et al.* 2009). Also, this approach has been utilized for the development of male-sterile transgenic plants using tapetum-specific promoters for the expression of the barnase gene (Gudynaite-Savitch *et al.* 2009). However, since each promoter responds differently to the activation potential of each particular enhancer, it appears that there is no single enhancer that can fully solve the problem (Singer *et al.* 2011).

In addition, inserting different DNA insulator sequences, mainly from non-plant species, between enhancers and nearby promoters in transgenic plants has also been suggested as a

strategy to block misexpression (Gudynaite-Savitch *et al.* 2009). Since it has been proven that plants do use insulators to regulate their gene expression, the main focus now is to identify, isolate, and characterize plant DNA insulators possessing enhancer-blocking activity. Finding such sequences would have a huge impact on the improvement of agricultural biotechnology by providing a novel way for reducing enhancer-promoter interactions during plant transformation experiments with composite vectors (Singer *et al.* 2011).

1.7 Hypotheses

This study aims to identify and characterize DNA insulators with enhancer-blocking activity in the model plant *Arabidopsis thaliana.* Two hypotheses were proposed:

1. Short DNA sequences and their binding proteins function as insulators in *A. thaliana*. These insulators will share sequence or mechanistic similarities with known insulators from other species;
2. DNA sequences with insulator activity can be identified from DNA libraries via a genetic screen of *A. thaliana*.

Chapter 2: Experimental Procedures

2.1 Plasmid construction

All constructs used in this project for plant transformation were derived from pCAMBIA series of plasmids described at http://www.cambia.org/daisy/cambia/materials/vectors.html. The base vectors used in the overall insulator project were constructed by Dr. L. Gudynaite-Savitch.pC1 was used for the initial selection of candidate insulator sequences. Candidate sequences were then cloned into pB31 to allow for GUS screening. An overview of their construction and use is provided in the Results section 3.1.

2.2 Molecular biology techniques

2.2.1 DNA isolation

Plasmid DNA for sub-cloning and sequencing was isolated from *Escherichia coli* DH5α (Life Technologies Inc. Burlington, Canada) using the GeneElute Plasmid Mini-Prep Kit (Sigma-Aldrich) according to the manufacturer's instructions direction except that the elution buffer was 0.1mM TrisHCl-0.01mM EDTA, pH8.0.

Plant DNA for PCR was isolated from rosette leaves of *Arabidopsis thaliana* by either using DNeasy Plant Kit (Qiagen) according to the manufacturer's recommendation or the DNA extraction method described by McKinney *et al*. 1995 with minor modifications.
PCR products were purified using the MinElute PCR purification kit (Qiagen) or after gel electrophoresis using the MinElute Gel extraction kit (Qiagen).

2.2.2 PCR amplification

A standard 20µLPCR reaction containing 10µL ddH_2O, Taq reaction buffer (10mM KCl, 10mM $(NH_4)_2SO_4$, 20mM Tris-HCl (pH 8.75), 0.1%Triton X-100, 0.1mg/ml BSA,2mM $MgSO_4$, 200nM each dNTP, 200nM forward primer, 200nMreverse primer, 1UTaq polymerase (UBI) and 1µL DNA template. A standard PCR program for amplification consisted of an initial denaturation step at 95°C for 5 minutes, followed by31 cycles of 30 seconds each of 94°C denaturation, 55°C annealing, and 72°C extension temperatures, ending with 72°C for 5 additional minutes. The best annealing temperature was determined experimentally by temperature gradient PCR for each pair of primers with a temperature range between (47.4°C-63.8°C). To produce sufficient DNA for further manipulation, such as cloning, the same PCR cocktail was scaled up to 100µL. Positive and negative PCR controls were included.

All PCR samples were analyzed by gel electrophoresis on 1.5% agarose gels in 0.5 x TBE (10 x TBE is 1M Tris base, 0.9M Boric acid, 0.01M EDTA), and run either at 89V for 60 minutes or 99V for 45 minutes. A 100bp DNA ladder was included a size standard.

2.2.3 DNA cloning

DNA fragments from transgenic plants were amplified by PCR using primers (**Table 2**) in order to introduce suitable restriction sites.

For sequencing, purified PCR products were cloned into the pGEM-T Easy vector (Life Technologies Inc. Burlington, Canada), and the presence of an insert was verified by PCR. Plasmid DNA was purified and sequenced (PBI, Saskatoon).

For sub-cloning of putative insulators into pB31, amplified PCR products were digested with the appropriate restriction enzyme (or in some cases two enzymes) and ligated into pB31 digested with the same enzyme(s). The presence of an insert was verified by PCR. Plasmid DNA was purified and sequenced.

A typical protocol follows. PCR fragments generated with the CLO- Eco For and CLO-Eco Rev primers were digested in a reaction with 5U EcoR1 (Fermentas), 2.5µL purified insert, 3.5µL water, and 2X Tango buffer (Fermentas) for 3 hours at 37^{o}C. Following heat inactivation (10 minutes at 60°C) the DNA was ligated into pB31 which had been digested with the same enzyme(s) using reagents provided by Promega: 1µL 10x Rapid Ligation Buffer, 50ng restricted pB31 and 3U T4 DNA ligase, 7µL of restricted PCR fragment. The ligation mixture was incubated overnight at room temperature.

2.2.4 Bacterial transformation

For the isolation of *E. coli* transformants, plasmid constructs carrying a DNA fragment of interest were introduced into sub-cloning efficiency *E. coli* DH5α (Life Technologies Inc. Burlington, Canada) according to the manufacturer's protocol. Transformants were selected on 100µg/mL *Ampicillin* or 50µg/mL *Kanamycin*.

Plasmids DNA was introduced into *A*grobacterium *tumefaciens* strain GV3101 (pMP90) (Koncz and Schell 1986) by electroporation (Wang *et al*. 2006). Transformants were selected on 50µg/mL *Kanamycin* and 100µg/mL *Rifampicin*.

2.3 Plant techniques

2.3.1 Plant growth conditions

A. thaliana (ecotype Columbia) plants were grown in growth cabinet model E15 (Conviron) under the following long day conditions: 16h, 22°C, 100µmol $m^{-2}s^{-1}$photosynthetically active radiation (PAR) day / 8h, 22°C night, relative humidity of 50%. The plants were watered approximately every two days with efforts made to keep the bottoms of the pots continuously wet. Hoagland's nutrient solution was used for watering once a week (Gudynaite-Savitch *et al*. 2007).

Rosette leaves from each transgenic line were collected in 1.5mL tubes and stored in a freezer at -20°C for DNA isolation.

2.3.2 Plant transformation

Agrobacterium-mediated transformation of *Arabidopsis thaliana* (L.) Heynh. (ecotype Columbia) was performed by the floral dip method (Clough and Bent 1998). Transformed plants were selected on half-strength Murashige and Skoog medium (Murashige and Skoog 1962) supplemented with 30µg/ml *Hygromycin-B* and 250µg/ml *Timentin*; (h*ygromycin-B* for selection of plants carrying the plasmid vectors, and *timentin* to inhibit bacterial growth). After 2-3 weeks on selective media, transgenic plants were transferred to soil and moved to a growth cabinet.

An inoculum was prepared by growing an *Agrobacterium* overnight culture in L-broth supplemented with *Rifampicin* and *Kanamycin* as above. The next day, this starter culture was diluted in 500mL of L broth, and grown overnight at 28°C with shaking. Bacteria were collected by centrifugation for 10 minutes at room temperature at 5000rpm

in a GSA rotor. The supernatant was removed and the cell pellet was suspended in an equal volume of a freshly made solution containing 25gm sucrose, 1.1g MS basal salts (Sigma) and 250μL Silwet L-77 in a volume of 500mL.

Plants inoculated with *Agrobacterium* were returned to the growth cabinet and watered until the plant reached maturity and was ready for harvesting. Seeds were then harvested and stored at 4^{o}C.

2.3.3 Staining of transgenic plants for GUS expression

Transgenic seeds collected from transgenic plants carrying a specific CLO insert such as, CLO I-3 were sprouted on selective plates with 30μg/ml *Hygromycin-B* and 250μg/ml *Timentin* in order to select for the transgenic plants carrying the transformed plasmid vectors. The plates were then placed in the growth cabinet for 10-12 days until visible roots and leaves formed. Each of these transgenic plants carrying the same insert was then moved into soil in separate pots to complete their growth in the growth chamber until enough leaves, flowers and siliques could be collected for GUS staining. For example, if three transgenic plants grew on the selective plates, the three were transferred into 3 different pots (pot #1, pot #2, and pot #3) and these transgenic plants were named I-3 #1, I-3 #2, and I-3 #3, respectively. Then one flower sample, which contained at least three stems with 5-6 flowers, was prepared for GUS staining while two leaf samples contain 3-4 leaves each, and two silique samples with 3-4 siliques each collected at different developmental stages, were prepared for the GUS staining. For example, from the I-3 #1 transgenic plant, one flower sample, and two leaf and two siliques samples were analyzed. The GUS histochemical staining was carried out as described by Malik *et*

al. (2002). Leaves and siliques were stained at two different times; the first when the plants were 4-5 weeks old and the second 7-10 days later.

The intensity of GUS staining was determined visually using the following scale: 0 (none) 1 (weak), 2 (medium), 3 (high) (see Supplementary Figure S1 of Gudynaite-Savitch *et al*. 2007) and if only one leaf or silique or stem or flower in the sample was positive, the entire sample was scored as positive.

The GUS histochemical staining was performed by using X-Gluc (5-bromo-4-chloro-3-indolyl-*β*-D-glucuroniccyclohexylammonium salt). Two different X-Gluc solutions were prepared; one for staining flowers and leaves, and one for the siliques. The flower and leaf X-Gluc stock solution was made by adding 57.7gm/L NaH_2PO_4 (H_2O), 42.3gm/L Na_2HPO_4 ($7H_2O$), 20gm/L EDTA (pH 8.0), 10gm/L $K_3Fe(CN)_6$, 10gm/L $K_4Fe(CN)_6$, and 1g/L Triton X-100 to 859mL water. The X-Gluc stock buffer for siliques was the same , except that the Triton X-100 concentration was increased to 2.5 g/L. 47mL of the stock solution was poured into a 50mL plastic tube then 0.5mL of both $K_3Fe(CN)_6$, and $K_4Fe(CN)_6$ was added to the solution. 50mg of X-Gluc dissolved in 50-100mL DMSO, was then added to this 50mL solution. The solution was stored at 4°C for frequent use, and wrapped in foil before storage.

Transgenic plant tissues were submerged in the X-Gluc buffer in labeled 1.5mL tube; 3-4 samples of the same tissue from a single plant were placed in the same tube. For silique staining, the siliques were cut length-wise on a Petri dish to allow entry of the stain. The plant samples in tubes were vacuum infiltrated for 30-40 minutes by setting the lab vacuum at 4bar. After removing the samples from the vacuum they were incubated overnight at 37°C. Then the samples were soaked in ethanol solutions with different

concentrations to bleach chlorophyll from the tissues. First, plant tissues were steeped in 30% ethanol for 2 hours, then soaked in 50% ethanol for an additional 2 hours, and finally, submerged in 70% ethanol. This last step was repeated as needed until the green pigment was no longer present. Stained tissues then were stored in 70% ethanol at room temperature.

2.3.4 PCR screening of transgenic plants

Each DNA sample from transgenic *A. thaliana* leaves was tested with three rounds of PCR to confirm DNA extraction and confirm that the plant had been transformed. The first round used SALK_049131_RP2 and SALK_049131_LP2 primers (**Table 2**) to amplify an *A. thaliana* basic helix-loop-helix DNA-binding protein gene (AT1G31050.1); the presence of this sequence confirmed successful DNA extraction. The second round used GUS-For/GUS-Rev primers (**Table 2**) to amplify the *GUS* gene found in the transgene encoding *β-glucuronidase*; the presence of this sequence indicated successful transformation and insertion of the transgenic construct into the *A. thaliana* genomic DNA. The third round used 35S46-For/35S46-Rev primers (**Table 2**) to amplify the 35S46 core promoter; the presence of this sequence also confirms that the transgenic construct was successfully inserted into the *A. thaliana* genomic DNA. Standard PCR conditions were used for all reactions.

All PCR samples were analyzed by gel electrophoresis. For screening for SALK and *GUS* genes 0.8% agarose gels were used. Screening for *35S* core promoter was done by using 1.25% agarose gel in 0.5x TBE. The predicted sizes of the PCR fragments were: the *GUS* gene ~2000bp; SALK ~700bp; and *35S* core promoter ~100bp. Putative

insulator sequences were amplified and identified by PCR using 1300LacZ-For/35S46-Rev primers (**Table 2**). The expected sizes ranged from 150 to 450bp, depending upon the individual potential insulator.

For sequencing, putative insulator sequences were amplified by PCR using 1300LacZ-For/ GUS 5'- Rev primers, purified and sent to PBI for sequencing with the 35S46-Rev primer.

2.4 Bioinformatics and sequence analysis

Routine analysis of DNA sequences was done with DNAMAN software (Lynnon Corporation). Webcutter 2.0 web site (http://rna.lundberg.gu.se/cutter2/) was used to identify and map all of the internal and external *Eco*RI and *Bam*HI restriction sites.

Primer	Target of Amplification	Sequence
SALK_ 049131_RP2	Genomic At DNA, positive control	5’-GTCTCTACCGTACGCGCTTC-3 (20mer)
SALK_ 049131_LP2	Genomic At DNA, positive control	5’-GGTTTGCATTTGACCTTTCG-3’ (20mer)
GUS- For	GUS gene in transgene	5'-CGTCCTGTAGAAACCCCAAC-3' (20mer)
GUS- Rev	GUS gene in transgene	5'-GTGGCTAGCTTGTTTGCCTC-3' (20 mer)
GUS 5’- Rev	CLO DNA insert, PCR for sequencing	5'-CTACAGGACGGACGAGTCGTC-3' (21 mer)
35S46-For	35S core promoter in transgene	5'-CTTAAGCTTCGCAAGACCCTTCCTCTATATAAG (33 mer)
35S46 –Rev	35S core promoter in transgene. Sequencing primer.	5'-CTACCATGGTCAAGAGTCCCC (21 mer)
1300LacZ- For	CLO DNA transgene insert	5'-CACTCATTAGGCACCCCAGG3 (20 mer)
CLO-Eco For	CLO DNA transgene insert; add restriction sites	5’-CGTTCTAGAGGATCCCGAATTCGAGACAAGC-3' (31mer)
CLO-Eco Rev	CLO DNA transgene insert; add restriction sites	5'-GCAGCTATTCCAAACCGCGGGAAAGATC-3’ (28mer)
CLO-For PHB	CLO DNA transgene insert; add restriction sites	5'-CGTCTGCAGAAGCTTGGATCCGAGACAAGC-3' (31mer)
CLO-Rev PHE	CLO DNA transgene insert; add restriction sites	5-'AGTCTGCAGAAGCTTGAATTCAGGAGGAGG-3' (30mer)

Table 2: Sequences of all primers used, with the target sequences for amplification and positive controls indicated.

Chapter 3: Results

3.1 Overview of the system for the isolation of insulator sequences from *Arabidopsis thaliana*

The system used to identify possible insulators was developed by Dr. L. Gudynaite-Savitch a post-doctoral fellow working in our lab and is unpublished. The system uses three different transformation vectors, pC1, pB31 and pL1, all derived from common pCAMBIA vectors (http://www.cambia.org/daisy/cambia/585.html) and expressing the gene *hptII* (*hygromycin phosphotransferase II*) encoding hygromycin resistance, driven by the CaMV*35S* enhancer/promoter system. In order to better understand this thesis I am providing a summary of the key elements in this system, how it was constructed and how it is used.

pC1 (pCAMBIA1300- 35S46-CodA) (**Figure 6**) is a negative selection vector. The conditional negative selective marker gene, *codA* (cytosine deaminase) is expressed from the core CaMV*35S* promoter (35S46). The level of expression from 35S46 is low; however, the enhancer in CaMV*35S* driving *hptII* can boost expression of c*odA*. Cytosine deaminase converts 5-FC on selective plates into 5-FU (5-Fluorouracil, **Figure 7)**, which is a toxic, leading to cell death. In plants and mammals enzymatic conversion of cytosine to uracil is absent; therefore, 5-FC does not affect the growth of wild-type plants at the levels used, while transgenic plants expressing cytosine deaminase are negatively selected on 5-FC plates. The presence of insulators between CaMV*35S*and 35S46 will block the enhancer and the plants will survive. DNA fragments were inserted upstream from the *35S* core promoter

in order to test their ability to block the influence of the *35S* enhancer/promoter on this promoter.

pB31 (pCAMBIA1300-35S 46-GUS, **Figure 8**) is a vector in which *codA* in pC1 has been replaced by GUS while the organization of the key remaining sequences is the same. Expression of GUS requires that the enhancer in CaMV*35S*influence 35S46 and once again the presence of insulators between CaMV*35S* and 35S46 will block the enhancer. If the sequence is a non-functional insulator all tissues will stain blue reflecting constitutive expression of the CaMV*35S* promoter. In the presence of a functional insulator we would expect no GUS expression. This vector is used to confirm that the sequence has insulator activity.

pL1(pCAMBIA 1391 Napin, Gudynaite-Savitch *et al*. 2009) extends the strategy of pB31 except that GUS expression is controlled by the seed-specific napin promoter that replaces 35S46 and like 35S46 can be influenced by the CaMV*35S* enhancer. If the potential insulator sequence is a non-functional, all tissues will stain blue but in the presence of a functional insulator we would expect GUS expression only in the seed. This vector is used to verify that the sequence has insulator activity with a second promoter. This vector was not used in the present project.

3.1.1 Statement of contributions

All pCAMBIA plasmid vectors used in this study were previously constructed by Dr. Loreta Gudynaite-Savitch, and others. Oligonucleotide library was synthesized by a machine outside the lab. 100 DNA fragments from the oligonucleotide library were first screened in the first selection system pC1 by Dr. Loreta Gudynaite-Savitch and Tatiana Semiz. 60 DNA

sequences were identified from the first screening of these fragments in pC1 and only the best 20 of them, which were selected based on their growth strength in the first selective plates with 5-FC, were re-cloned into the second selection system pB31. 17/20 of these potential insulator sequences were cloned into pB31 by Dr. Loreta Gudynaite-Savitch. The remaining three potential insulator sequences (CLO III-53, CLO III-58, and CLO III-74) were cloned into the pB31 by me. Dr. Douglas Johnson was screened some of these cloned sequences in pB31 for possessing enhancer-blocking activity using the histochemical GUS staining. For some clones the number of samples were not sufficient, thus, I re-transferred and re-stained them in order to again adequate number of samples for each cloned. For the other cloned that had not transferred before, I transferred and stained them. Also, for the promising candidate that was identified from the GUS staining, CLO I-3, the first and the second replicates were transformed into plants by the floral dip method and stained by me. Furthermore, I extracted DNA samples from X transgenic plants leaves carrying the insert I-3 and screened them by PCR to confirm the presence of certain sequences including; the insert I-3, GUS gene, and 35S46 core promoter. Then I generated PCR fragments ready for cloning, and a summer student, Dora Laczko, helped to finish this part of the project by preparing the insert I-3 PCR products to be ready for sequencing.

3.2 Identification of potential insulator sequences

Two different libraries had been constructed in pC1 (**Figure 6**). As an initial step toward the discovery of a novel plant DNA insulator, a random oligonucleotide library was generated by DNA synthesis, converted to dsDNA by PCR and cloned into pC1 using BamH1 and EcoR1 restriction/ligation to give pCLib (**Figure 6**). This library consists of

154bp fragments in which 124bp of random sequence are flanked by 30bp of defined sequences that act as sites for PCR primers and contain BamHI and EcoRI sites for cloning. A second library was generated from total *A. thaliana* DNA by restriction digestion, followed by size selection by gel electrophoresis and cloning. This library has not been screened.

Initially 60 transgenic plants that survived selection were obtained; these may contain fragments that possess insulator activity. Total DNA was isolated and vector inserts were PCR amplified, sequenced, and classified into 4 groups (I, II, III, and IV based upon the date of selection). The best 20 sequences out of the 60 sequences that were identified from the first selection system pC1, were chosen based on their growth strength on the selective plates with 5-FC for re-cloning (individually) into a second selection plasmid vector pB31.17 sequences out of these 20 putative DNA insulator sequences were cloned into pB31 plasmid vectors by Gudynaite-Savitch while the three remaining sequences (CLO III-53, CLO III-58, and CLO III-74) were cloned during this project. Cloning examples are shown in **Figure 9**. Because of intellectual property issues, sequences are not presented.

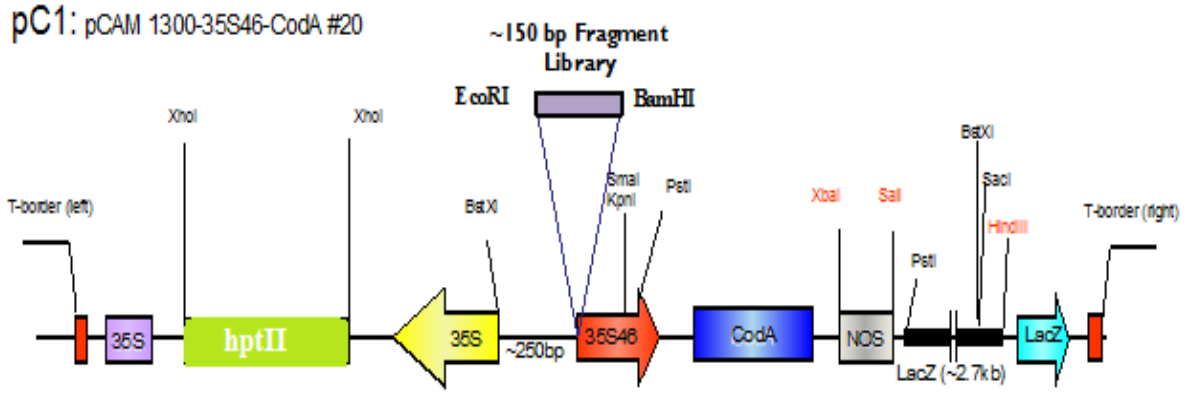


Figure 6

Schematic diagram of the pC1vector (pCAM 1300 35S46 CodA) indicating the position used to clone potential insulator sequences.

Features:
yellow arrow – CaMV*35S* enhancer, driving constitutive expression of the green *hptII* reporter gene encoding *hygromycin* resistance (with light purple 35S terminator sequence);
red arrow – CaMV 35S46 core promoter, poorly functional without activation by *35S* enhancer, driving expression of blue *codA* reporter gene encoding cytosine deaminase enzyme (with grey NOS terminator sequence);
purple bar – insertion site of the random oligonucleotide fragments
black – lacZ gene from yeast
red – T-DNA L/R borders delineating the borders of the transgenic insert.

Overall mechanism of function: if the light grey oligonucleotide DNA insert functions as an insulator, the yellow *35S* enhancer will not interact with the red 35S46 core promoter and the plant will not express *codA*'s product; this stops a metabolic pathway given in **Figure 7** and allows normal plant growth.

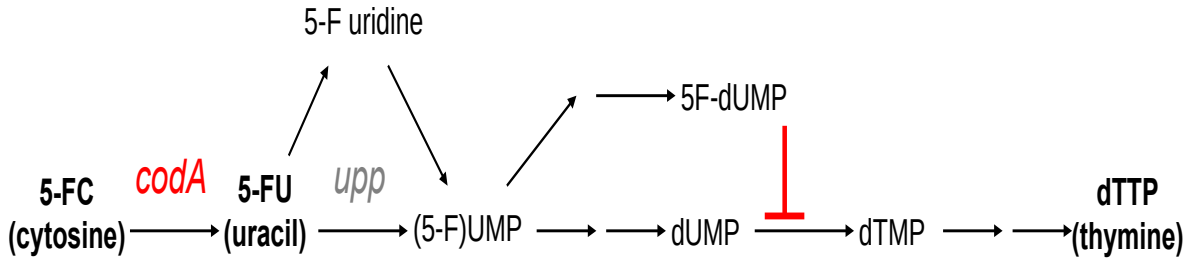


Figure 7

The toxic metabolic pathway of 5-FC catalyzed by the product of *codA* expression. Normally cytosine is metabolized to uracil, and eventually to dTTP (deoxythymine triphosphate) that is required for all DNA synthesis. Cytosine deaminase enzyme encoded by *codA* converts5-FC to 5-FU, which is then processed to 5F-dUMP (5-fluoro-deoxyuridine monophosphate) indirectly, via the intermediate 5F-uridine or directly, by uracil phosphoribosyltransferase (*upp*). 5F-dUMP irreversibly inhibits thymidylate synthase activity, and as a result the cells are deprived of deoxythymidine triphosphate (dTTP) necessary for DNA synthesis.

Modified from (http://www.invivogen.com/MSDS/MSDS_5-FC.pdf).

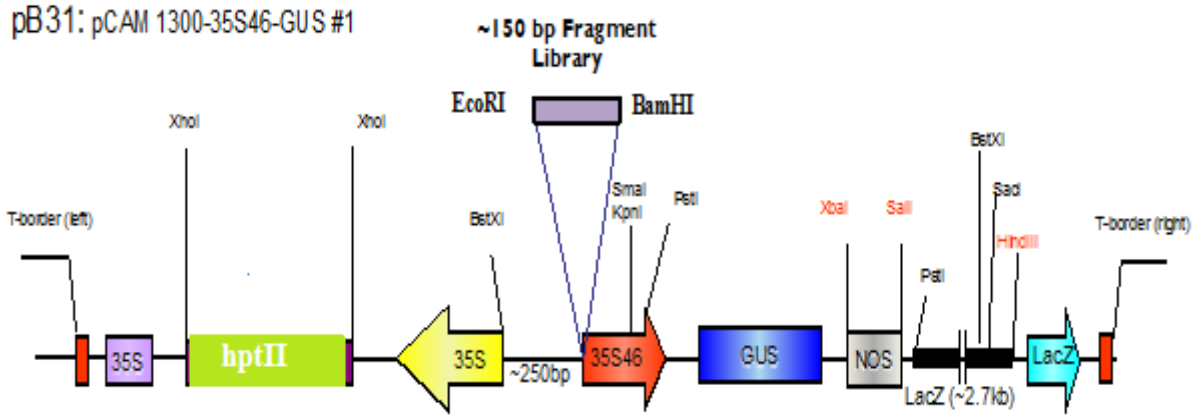


Figure 8

Schematic diagram of the pB31 vector system used to transform *A. thaliana* with a random oligonucleotide DNA fragments.

Features:
yellow arrow – CaMV*35S* enhancer, driving constitutive expression of the green *hptII* reporter gene encoding *hygromycin* resistance (with light purple 35S terminator sequence);
red arrow – CaMV*35S*46 core promoter, poorly functional without activation by *35S* enhancer, driving expression of blue *GUS* reporter gene encoding *β- glucuronidase* enzyme (with grey NOS terminator sequence);
purple bar – insertion site of the random oligonucleotide fragments
black – lacZ gene from yeast
red – T-DNA L/R borders delineating the borders of the transgenic insert.

Overall mechanism of function: if the light grey oligonucleotide DNA insert functions as an insulator, the yellow *35S* enhancer will not interact with the red *35S*46 core promoter and the plant will not express the blue *GUS* reporter gene.

The pB31 plasmid vectors with the correct inserts were transformed into *Agrobacterium tumefaciens* and then used to transform *A. thaliana* plants using the floral dip method. Seeds from transgenic plants were germinated in selective plates containing *Hygromycin-B* and *Timentin*. Surviving plants were transferred into soil to grow until *GUS* activity was assayed. The GUS assay is easy to perform, sensitive, relatively inexpensive, highly reliable, safe, and requires no specialized equipment, and is highly visual (Karcher 2002). Transgenic plant tissues were categorized based on the intensity of the GUS staining that was determined visually. **Table 3** summarizes the staining results while the staining patterns for individual transformant are given in the Appendix **Table A1**. Data for only 19 of the 20 clones is presented as it was found that two had the same sequence.

The results from **Table 3** can be summarized as follows:

For clones (CLO I-2, CLO III-58, and CLO III-80) insufficient data was collected to form a conclusion. These will need to be repeated. For clones (CLO II-3, CLO II-7, CLO III-4, CLO III-17, CLO III-22, CLO III-27, CLO III-53, CLO III-55, CLO III-57, CLO III-63, and CLO III-74) there appears to be no insulator activity. Some of these have sufficiently high numbers of samples that were analysed but in other cases (CLO III-63, and CLO III-74) the number of samples was low; however, the percentage that stained blue was high. For clones (CLO II-12, and CLO III-52) there appears to be weak insulator activity in some tissues but strong insulator activity in the siliques. While these candidates will need to be retested to obtain more transformed plants, they may represent possible tissue-specific insulator activity.

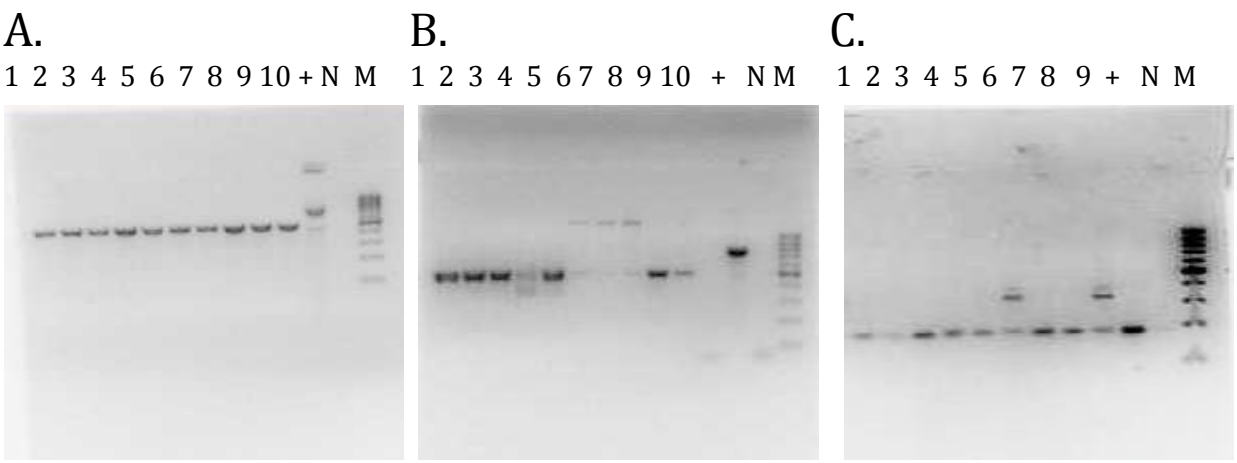


Figure 9

Agarose gel electrophoresis of colony PCR products from cloning of CLO fragments into pB31.

Following ligation and transformation, potential transformants were subjected to colony PCR. All PCR fragments were of the expected sizes.

A. CLO III-53 inserts from 10 colonies (lanes 1 to 10) with an expected size of ~500bp amplified with the CLO Eco-Rev and GUS 5'-Rev primers. CLO III-4 insert (expected size ~700bp) was used as positive control (+). N is negative control (water) and M is a 100bp ladder.

B. CLO III-74 inserts from 5 colonies (lanes 1-5) with an expected size ~ 500bp amplified with the CLO Eco-For and GUS 5'-Rev primers. Lanes 6-10 contained the same samples amplified by CLO Eco-Rev and GUS 5'-rev primers. CLO III-80 insert (expected size ~700bp) was used as positive control (+). N is negative control (water) and M is a 100bp ladder.

C. CLO III-58 inserts from 9 colonies (lanes 1-9) with an expected size ~180bp amplified with the CLO Eco-For and CLO Eco-Rev primers. CLO III-58 insert (expected size 150bp) was used as positive control (+). N is negative control (water) and M is a 100bp ladder.

Name	Insert size (bp)	Trans. events	Number of individual transformants of each clone	GUS Positive Flowers (Number)	GUS Positive Flowers (Percent)	GUS Positive Leaves (Number)	GUS Positive Leaves (Percent)	GUS Positive Siliques (Number)	GUS Positive Siliques (Percent)	Status	Classification
CLO I-2	446	Two	6	3	50.00%	0	0.00%	0	0.00%	Needs repetition	Insufficient data
CLO I-3	438	Two	49	0	0.00%	2	4.08%	0	0.00%	Finished	Candidate insulator
CLO II-3	151	Two	27	10	37.00%	19	70.30%	21	77.80%	Finished	Possible insulator activity
CLO II-7	150	Two	30	11	36.70%	21	70.00%	10	33.33%	Finished	No insulator activity
CLO II-10	154	Two	36	23	63.90%	13	36.11%	12	33.33%	Needs repetition	Possible insulator activity
CLO II-12	427	Two	15	5	33.30%	3	20.00%	0	0.00%	Needs repetition	Possible tissue-specific activity
CLO III-4	440	One	28	18	64.30%	4	14.29%	3	10.71%	Finished	No insulator activity
CLO III-17	443	One	30	25	83.30%	15	50.00%	2	6.67%	Finished	No insulator activity
CLO III-22	425	Two	36	26	72.20%	20	55.56%	3	8.33%	Finished	No insulator activity
CLO III-27	128	One	13	13	100.00%	5	38.46%	2	15.38%	Finished	No insulator activity
CLO III-52	446	Two	15	3	20.00%	2	13.33%	0	0.00%	Needs repetition	Possible tissue-specific activity
CLO III-53	152	One	15	13	86.70%	14	93.33%	14	93.33%	Finished	No insulator activity
CLO III-55	432	Two	23	23	100.00%	5	21.74%	2	8.70%	Finished	No insulator activity
CLO III-57	154	One	16	2	12.50%	13	81.25%	16	100.00%	Finished	No insulator activity
CLO III-58	150	One	3	0	0.00%	0	0.00%	0	0.00%	Needs repetition	Insufficient data
CLO III-63	154	One	7	2	28.60%	4	57.14%	7	100.00%	Finished	No insulator activity
CLO III-74	153	One	6	3	50.00%	6	100.00%	5	83.33%	Finished	No insulator activity
CLO III-78	450	Two	33	15	45.50%	6	18.18%	4	12.12%	Needs repetition	Possible insulator activity
CLO III-80	451	One	9	1	11.10%	0	0.00%	0	0.00%	Need more seed	Insufficient data

Table 3: The results of GUS staining with numbers and percentages in three different plant tissues (flower, leaf, silique) of CLO inserts in pB31 vectors. Orange = promising candidate CLO I-3; yellow = CLO sequences with possible insulator activity; green = CLO sequences with possible tissue-specific activity.*Silique samples include (silique + seeds). For details of the scoring system see Materials and Methods 2.3.3

Of the 19 clones that were tested only CLO I-3 with a 438bp insert has insulator activity in *A. thaliana*. This conclusion is based on the results of two independent transformation events in which a total of 49 transgenic plants were analysed for GUS staining. In these transgenic plants no flower or silique + seed samples stained blue while only 2 transgenic leaf samples stained blue indicating misexpression. Even in each of these two transgenic leaf samples, only a few leaves (1 out of 3 in each sample) actually stained blue and at the lowest level of detection ("1").

GUS staining was done once for some clones (**Table 3**); and no replicate was done for these clones because the GUS staining appeared with high percentages in all transgenic plant tissues (**Figure 10**) indicating that these sequences have no insulator activity. If a candidate looked promising (minimal or no GUS staining in all transgenic plant tissues in the first screen replicate) such as CLO I-3 (**Figure 10c1**), the GUS staining was repeated on transgenic plants from an independent transformation for confirmation of the results (**Table 3**). The stained transgenic plant tissues samples that were obtained from each potential insulator candidate were usually clear enough to decide if these sequences really might possess an enhancer-blocking activity; however in some cases the results from independent transformations were contradictory, for example for transgenic plant tissues containing CLO II-10 and CLO III-78.

For insert II-10 (**Figure 10c2**), the first screened selection had 15transgenic plants with no GUS staining in almost all transgenic plant tissues except for two flower samples with a faint GUS staining. Due to the death of many plants, GUS staining was not obtained from sufficient numbers of samples; thus a new transformation was done. However, in this replicate all flower had light GUS staining and most silique samples

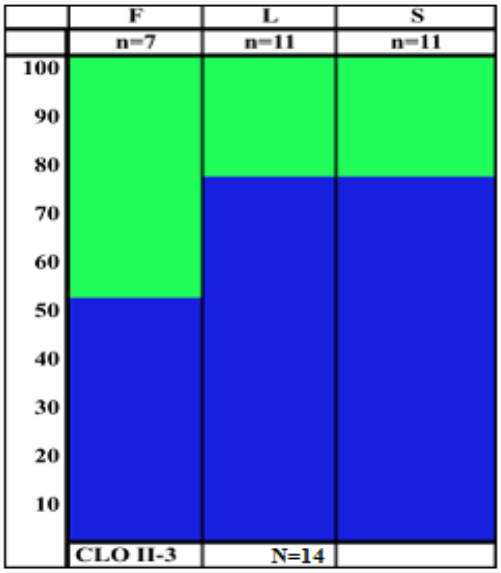
(12/21) had GUS staining ranging between medium to high (**Appendix Table A1)**. For insert CLOIII-78 (**Figure 10c2**), two transformations were done for a total of 33transgenic plant samples. All 17 samples stained from the first transformation had no GUS staining at all, whereas, the staining was observed in most transgenic flower samples (15/16) a few silique samples (**Appendix Table A1)** of the second transformation.

Further analysis of the GUS staining results has identified CLO sequences where specific transgenic plant tissues had no GUS staining at all while the other transgenic plant tissues had GUS staining which may imply that these sequences possess tissue-specific activity. For example, plants containing inserts CLO II-12 or CLO III-52 (**Figure 10b1&2**) had no silique staining whereas they had low flower and leaf staining.

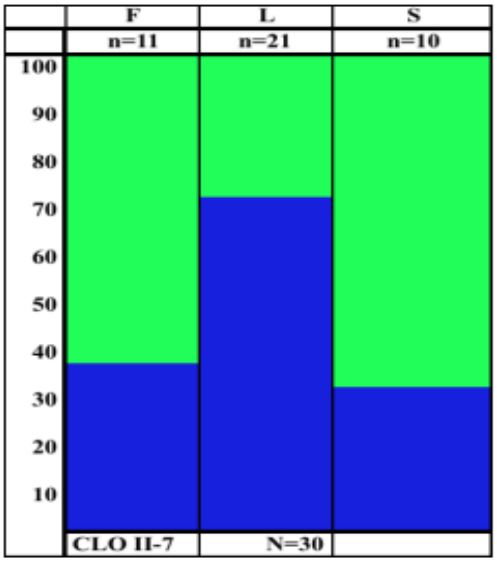
Thus for many of the CLO inserts described above the transformation experiments must be repeated to resolve the problems of low numbers and contradictory results between two independent transformation events; in addition confirmation of potential tissue-specific insulation must be confirmed.

a

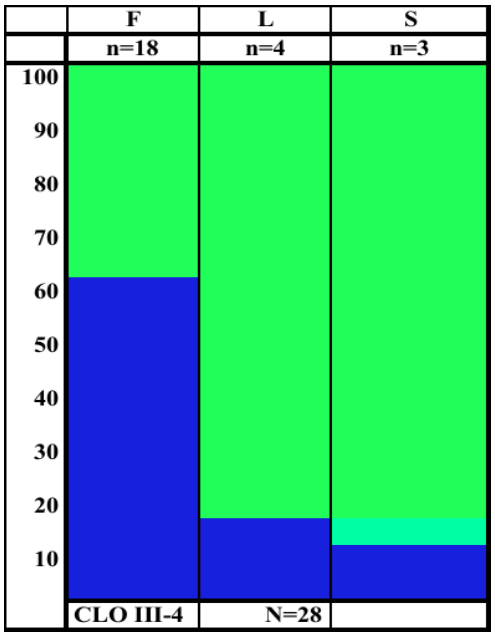
1


F
L
S
n=7
n=11
n=11
100
90
80
70
60
50
40
30
20
10
CLO II-3
N=14

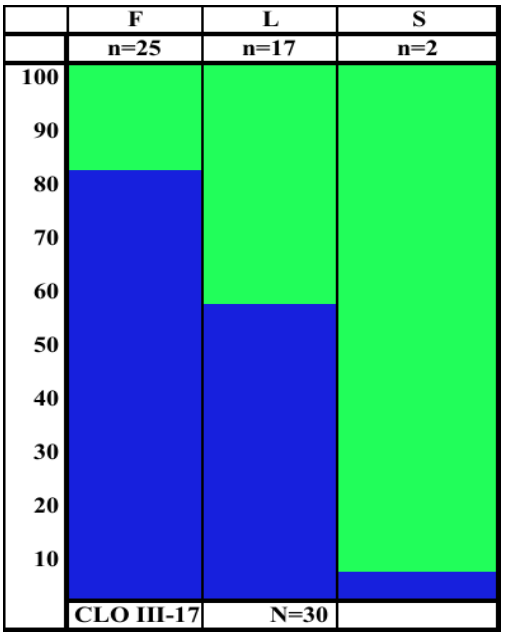
2


F
L
S
n=11
n=21
n=10
100
90
80
70
60
50
40
30
20
10
CLO II-7
N=30

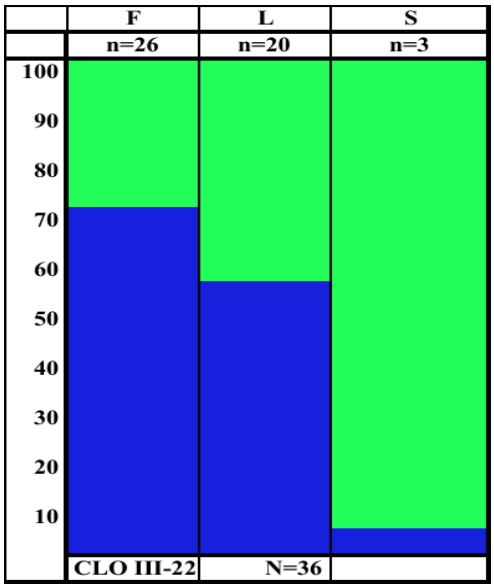
3


F
L
S
n=18
n=4
n=3
100
90
80
70
60
50
40
30
20
10
CLO III-4
N=28

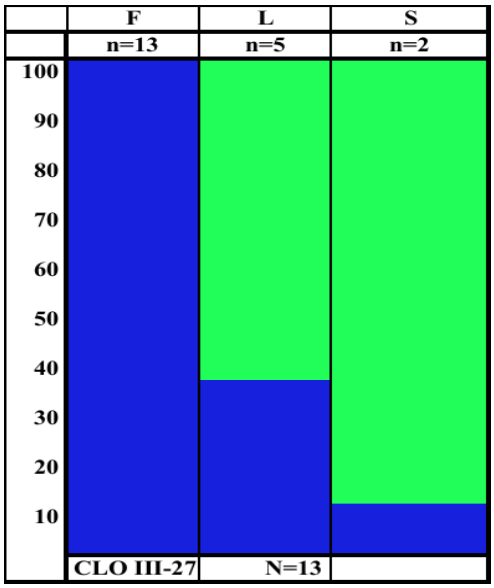
4


F
L
S
n=25
n=17
n=2
100
90
80
70
60
50
40
30
20
10
CLO III-17
N=30

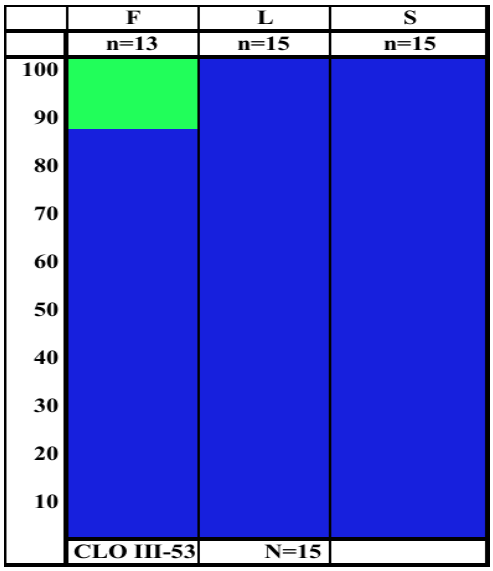
5


F
L
S
n=26
n=20
n=3
100
90
80
70
60
50
40
30
20
10
CLO III-22
N=36

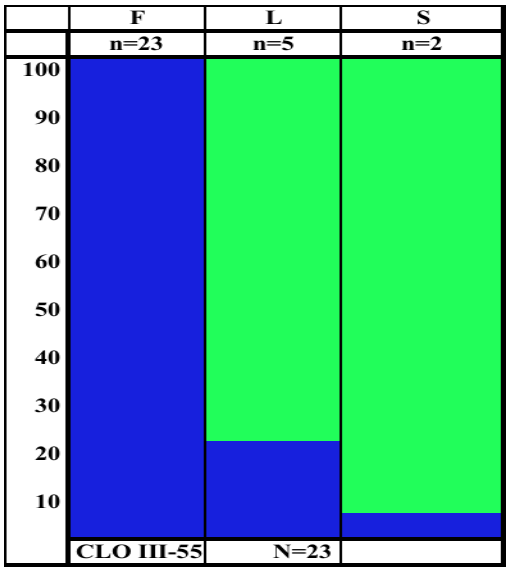
6


F
L
S
n=13
n=5
n=2
100
90
80
70
60
50
40
30
20
10
CLO III-27
N=13

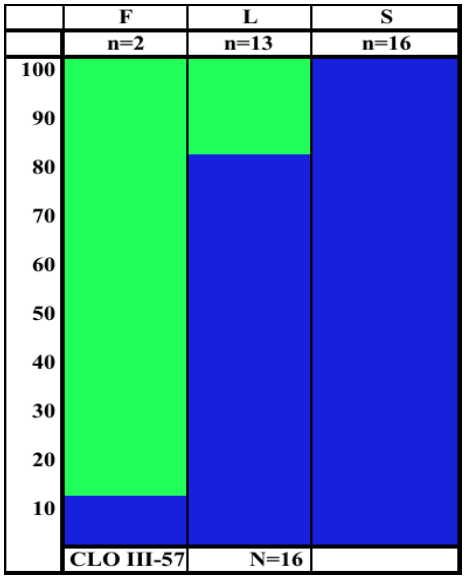
7


F
L
S
n=13
n=15
n=15
100
90
80
70
60
50
40
30
20
10
CLO III-53
N=15

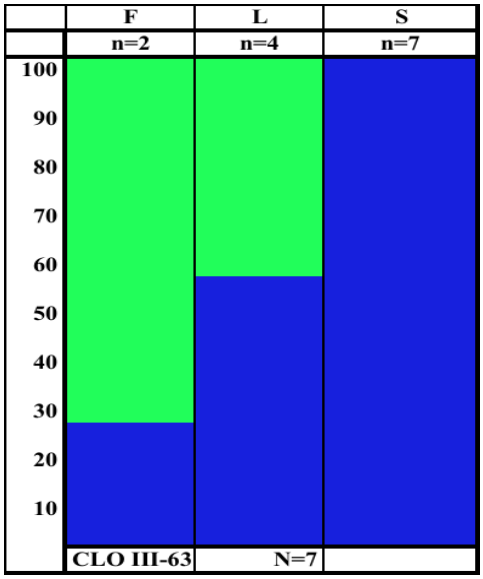
8


F
L
S
n=23
n=5
n=2
100
90
80
70
60
50
40
30
20
10
CLO III-55
N=23

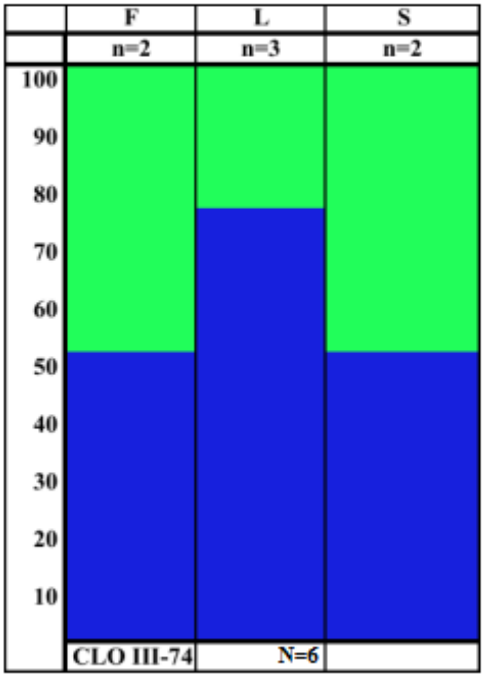
9


F
L
S
n=2
n=13
n=16
100
90
80
70
60
50
40
30
20
10
CLO III-57
N=16

10

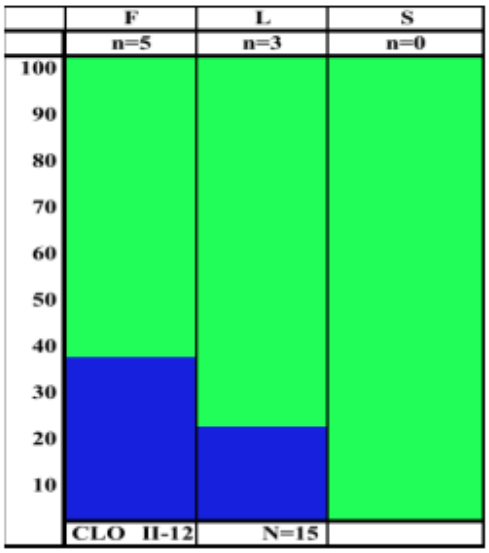

F
L
S
n=2
n=4
n=7
100
90
80
70
60
50
40
30
20
10
CLO III-63
N=7

11

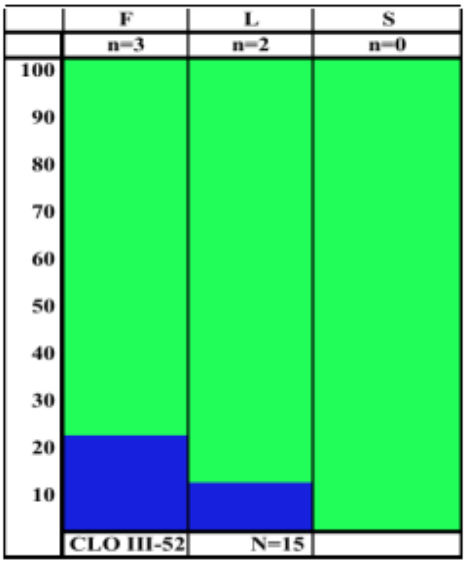

F
L
S
n=2
n=3
n=2
100
90
80
70
60
50
40
30
20
10
CLO III-74
N=6

b

1

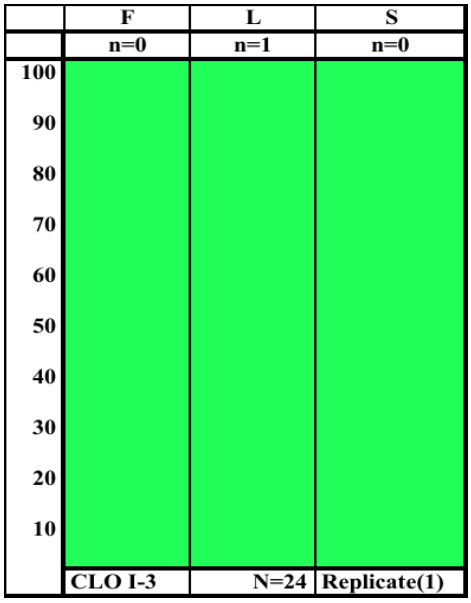

F
L
S
n=5
n=3
n=0
100
90
80
70
60
50
40
30
20
10
CLO II-12
N=15

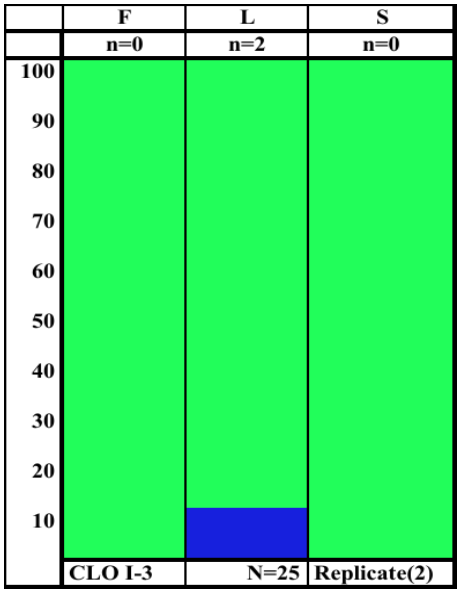
2


F
L
S
n=3
n=2
n=0
100
90
80
70
60
50
40
30
20
10
CLO III-52
N=15

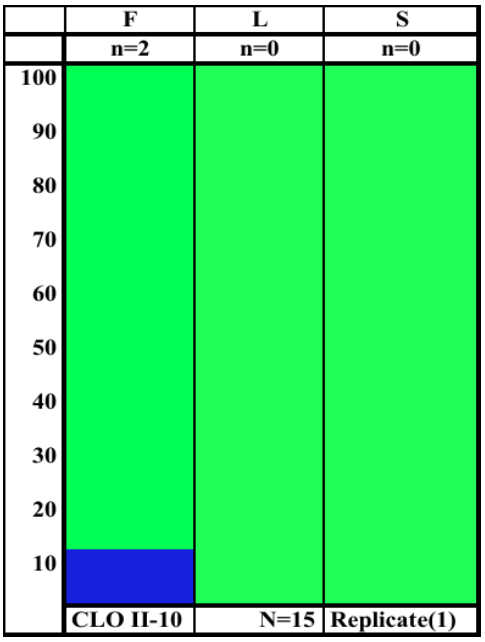
C

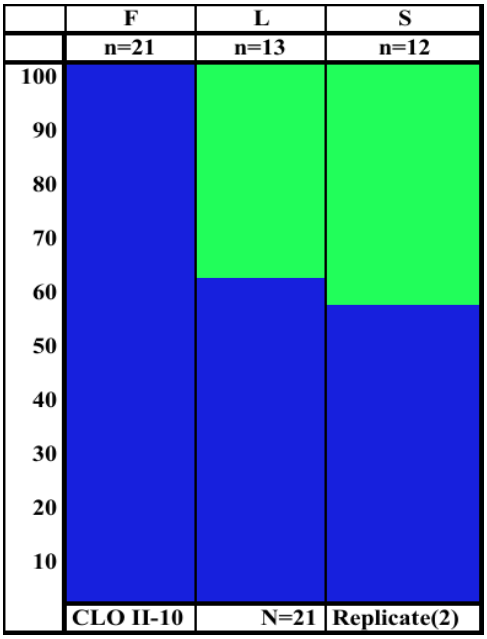
1


F
L
S
n=0
n=1
n=0
100
90
80
70
60
50
40
30
20
10
CLO I-3
N=24
Replicate(1)

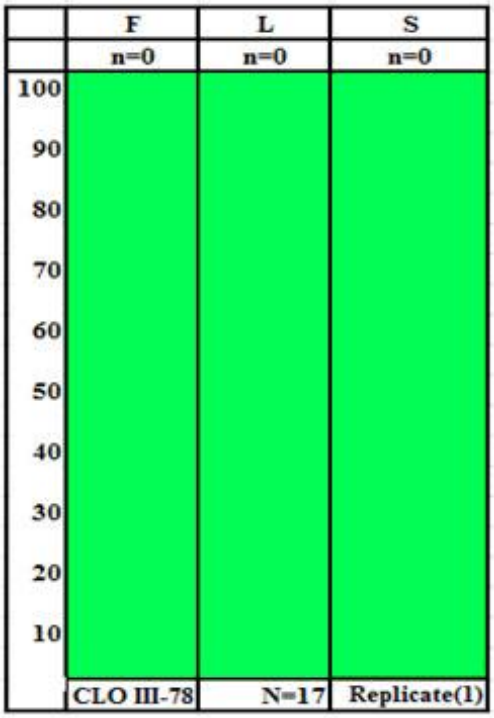

F
L
S
n=0
n=2
n=0
100
90
80
70
60
50
40
30
20
10
CLO I-3
N=25
Replicate(2)

2


F
L
S
n=2
n=0
n=0
100
90
80
70
60
50
40
30
20
10
CLO II-10
N=15
Replicate(1)


F
L
S
n=21
n=13
n=12
100
90
80
70
60
50
40
30
20
10
CLO II-10
N=21
Replicate(2)

3


F
L
S
n=0
n=0
n=0
100
90
80
70
60
50
40
30
20
10
CLO III-78
N=17
Replicate(1)

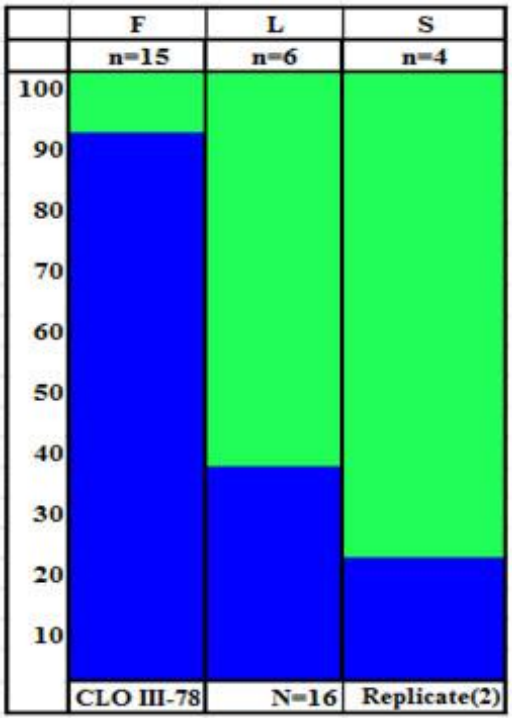

F
L
S
n=15
n=6
n=4
100
90
80
70
60
50
40
30
20
10
CLO III-78
N=16
Replicate(2)

Figure 10

Graphic representation of GUS staining results of the first and second replicates of each clone.

Leaves and siliques were stained at two different developmental stages and if only one leaf or silique or stem or flower in the sample was positive, the entire sample was scored as positive. See Materials and Methods 2.3.3 for further details.

Bar charts represent the percentages of GUS staining in each transgenic plant tissue (flower, leaf, silique) of selected CLO oligonucleotide sequences. The y-axis is the number of GUS stained samples as a percentage.
F= flower + stem; L= leaf; S= siliques + seed; n = number of samples with GUS staining in each transgenic tissues; N= total number of transgenic plant samples. Green = no GUS staining; blue = GUS staining

10a: Staining results for transgenic plant tissue samples with CLO inserts that appear not to have insulator activity.
CLO II-3, CLO II-7, CLO III-4, CLO III-17, CLO III-22, CLO III-27, CLO III-53, CLO III-55, CLO III-57, CLO III-63, CLO III-74.

10b: Staining results for transgenic plant tissue samples with CLO inserts that may have tissue-specific insulator activity.
CLO II-12, CLO III-52.

10c: Staining results for transgenic plant tissue samples with CLO inserts of ~450bp that may have insulator activity.
CLO I-3, CLO II-10, CLO III-78.

3.3 Analysis of CLO transformants

DNA extraction was performed on transgenic leaf samples collected from selected transgenic plants carrying the three promising oligonucleotide sequences (CLO I-3, CLO II-10, and CLO III-78), followed by a series of PCR amplifications to test for the presence of specific sequences. The first PCR screening, done using SALK_ 049131_RP2 and SALK_ 049131_LP2 primers (**Table 2**), which amplify a housekeeping gene present in the *A. thaliana* genome, tested the quality of the DNA preparation by the presence of a 750bp long DNA amplimer. A minimum of 12 samples was tested for each line with an example given in (**Figure 11**).

Additional PCR screening was performed to confirm the presence of the CLO fragment, the GUS transgene and CaMV35S46 promoter (data not shown). PCR using primers 1300lacZ-For and 3546-Rev which anneal to sites flanking the CLO insertion site (1300lacZ-For binding within the 250bp spacer region and 3546-Rev to the CaMV*35S*46 promoter, see **Figure 8**) amplified the CLO I-3, CLO II-10, and CLO III-74 inserts along with the flanking regions in pB31 (**Figure 12a**). Similarly, the primers GUS-For and GUS-Rev amplify the *GUS* gene (**Figure 12b**).

To confirm the identity of the PCR products for CLO I-3, amplimers were purified by agarose gel electrophoresis and sequenced directly or cloned into pGEM-T and sequenced. The results confirm that the transformants contained the expected CLO1-3 insert.

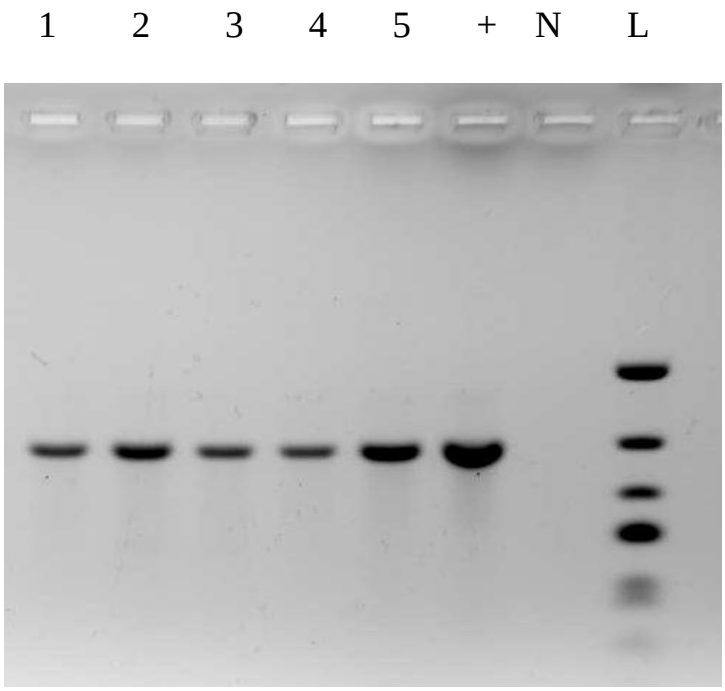


Figure 11

Example of agarose gel electrophoresis showing PCR amplification with SALK primers.
Transgenic *A. thaliana* DNA samples #1-5 containing the CLO I -3 insert are shown.
+ = positive control (wild-type *A. thaliana* DNA extract);
N= negative control (water);
L = pφX174HaeIII size ladder.

The expected size of ~750bp was observed in all cases.

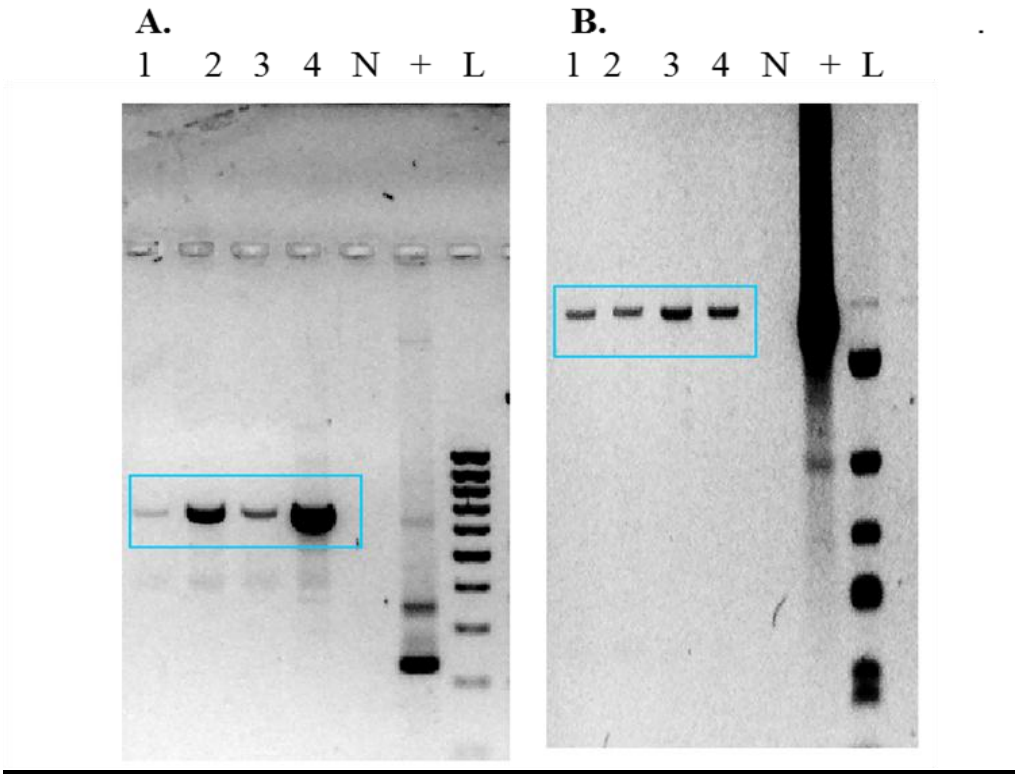


Figure 12

Example of agarose gel electrophoresis showing PCR amplification of DNA from CLO I-3 transgenic plants.

In both cases N is the negative control (water), + = positive control and L is the size ladder.

A. Amplification of the CLO 1-3 sequence and flanking pB31 sequences with the LacZ1300-For and 35S46-Rev primers. Samples #1-4 are DNA from CLO I-3 transgenic plan leaves with an expected size of ~638bp for the amplified fragments. The positive control is the CLO III-58 insert with an expected size of ~350bp for the amplified fragments. The size ladder is the Low Range DNA ladder (Fermentas SM0383).

B. Amplification of the *GUS* gene with the GUS-For and GUS-Rev primers. Samples #1-4 are DNA from CLO I-3 transgenic plant leaves. The positive control is pB31, with an expected size of ~2kb for the amplified fragments. The size ladder is pφX174 *Hae*III.

3.4 Bioinformatics and sequence analysis

The random oligonucleotide sequences as designed comprised 124bp fragments flanked by *Eco*RI and *Bam*HI sites for cloning. Sequencing results of the oligonucleotide sequences which were identified from the first selection system pC1 showed a large number of inserts with sizes of ~450bp long. To help to understand the origin of these fragments we aligned the long oligonucleotide sequences using the DNAMAN software, labeled all the internal and external *Eco*RI and *Bam*HI restriction sites using the Webcutter2.0 site (http://rna.lundberg.gu.se/cutter2/), and drew a map for restriction sites for each sequence to gain a better understanding of how such a phenomenon just happened (**Figure 14**). Mapping all of the *Eco*RI and *Bam*HI restriction sites (**Figure 13** and **Table 4)** indicted that in most cases three 150bp oligonucleotide sequences ligated to each other at their *Eco*RI and *Bam*HI flanking sites generating a long sequence consisting of ~450bp. Moreover, each 450bp oligonucleotide sequence was checked to see if there were repeats of the same sequence within it but all appear to contain three different sequences. Thus we proposed a model (**Figure 14**) in which 3 different ~150bp oligonucleotide sequences form a ~450bp fragment prior to ligation into pC1. As **Figure 13** suggests and sequence analysis confirms, the DNA sequence next to the restriction site contains PCR primer sites synthesized in the input oligonucleotide.

All of the 19 CLO oligonucleotide sequences cloned into pB31wereanalyzed to see whether they contained CCCTC, a sequence implicated in insulator activity in several species (**Table 1**). The majority of the 19 CLO oligonucleotides contain at least one copy of the short motif CCCTC (**Table 5**), suggesting that this sequence alone does not function as an insulator in *Arabidopsis.*

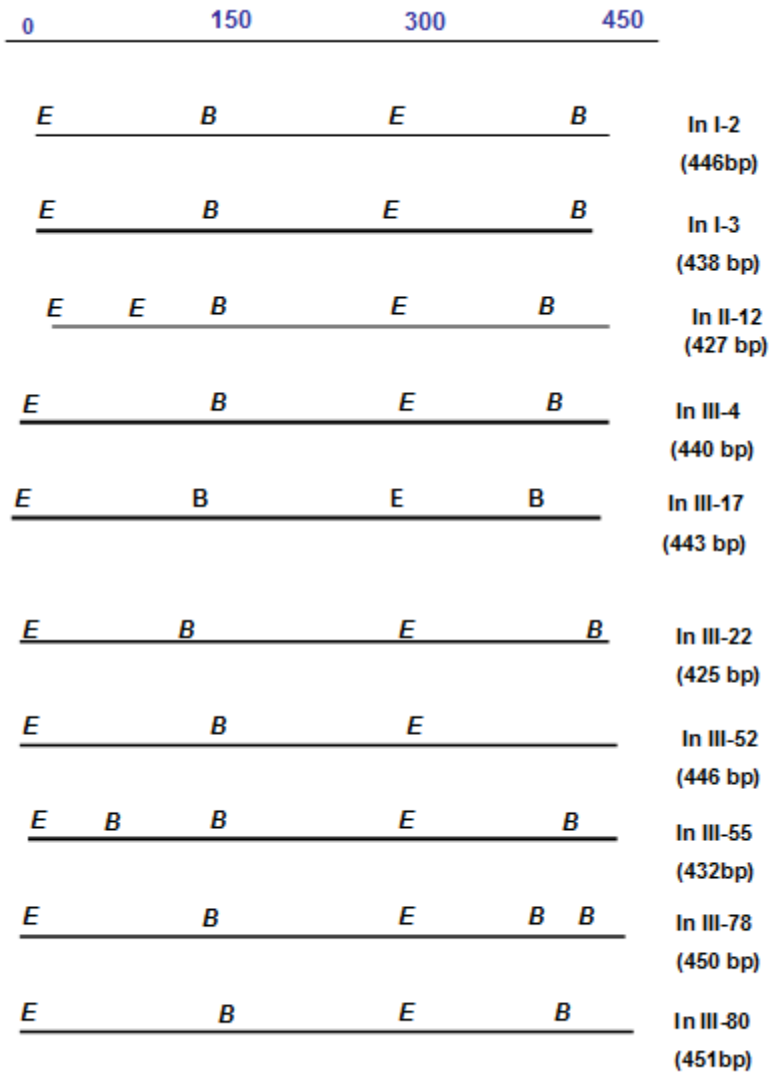


Figure 13

Maps of *Eco*RI and *Bam*HI restriction sites present on the long oligonucleotide sequences. First line shows the scale used to locate *Eco*RI and *Bam*HI restriction sites on each insert. E stands for *Eco*RI restriction site and B stands for *Bam*HI restriction site.

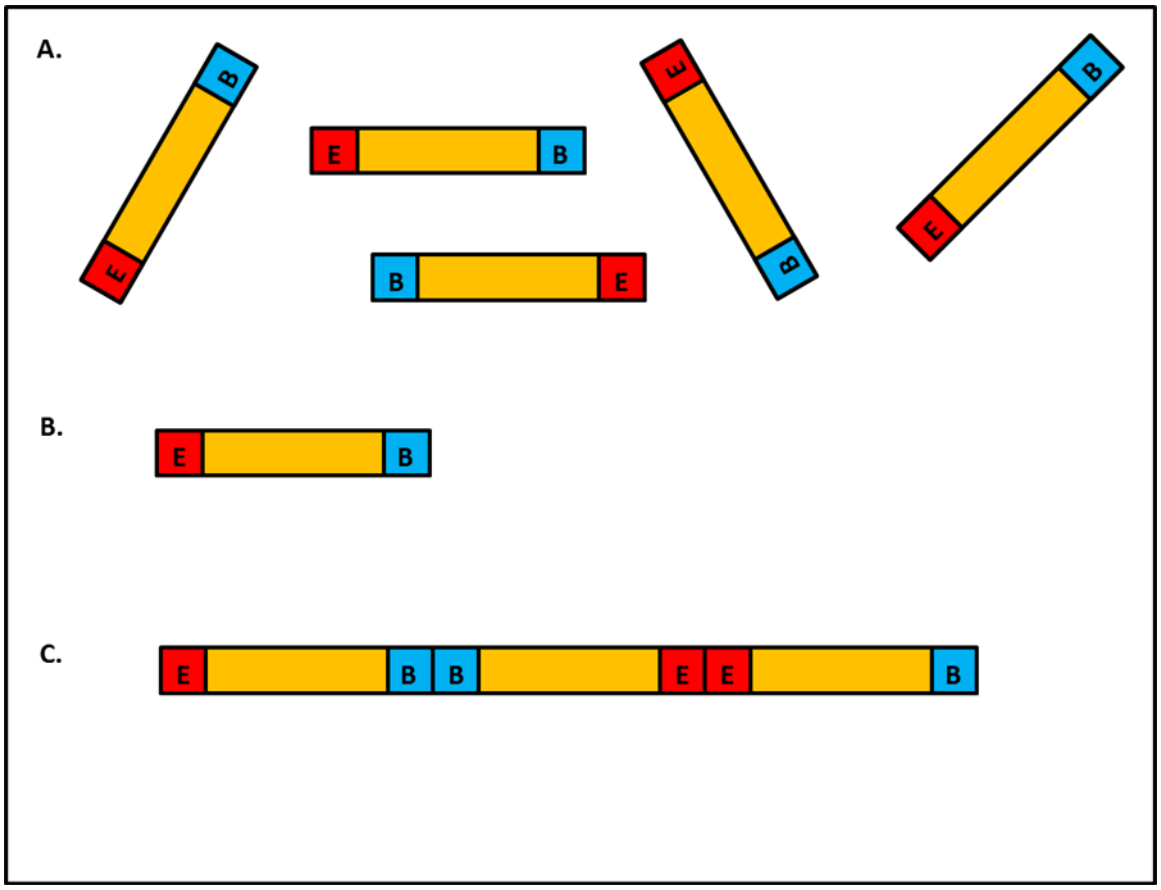


Figure 14

Model explaining the possible ligation pathways for 150bp oligonucleotide sequences.
Most the 450bp fragments consist of three 150bp oligonucleotide sequences ligated to each other at the EcoRI (E.) and BamHI (B.) sites marking the ends of each sequence. BB and EE indicate the point of ligation of two have sites into a single restriction site.

A. Pool of PCR generated fragments

B. Cloning of a single fragment

C. Cloning of three fragments

In B. and C. the vector pB31 is not shown.

CLO insert (size)	Location of *Eco*RI sites	Location of *Bam*HI sites
CLO I-2 (446bp)	1, 298	146, 441
CLO I-3 (438bp)	1, 297	148, 433
CLO II-12 (472bp)	1, 45, 287	139, 422
CLO III-4 (440bp)	1, 291	149, 435
CLO III-17 (443bp)	1, 291	146, 438
CLO III-22 (425bp)	1, 272	136, 420
CLO III-52 (446bp)	1, 295	148
CLO III-55 (432bp)	1, 282	63, 143, 427
CLO III-78 (450bp)	1, 281	137, 428, 447
CLO III-80 (451bp)	1, 297	149, 446

Table 4: The location of *Eco*RI and *Bam*HI restriction sites in the long oligonucleotide sequences cloned into pB31 vectors.

Inserts in pB31 vector	Insert size (bp)	The copy number of the CCCTC sequence in each insert	The location of the CCCTC sequence on each insert
CLO I-2	446	3	116-120,321-325, 415-419
CLO I-3	438	2	58-62, 336-340
CLO II-3	151	1	76-80
CLO II-10	154	3	34-38, 111-115, 135-139
CLO II-12	427	3	79-83, 312-317, 370-374
CLO III-4	440	2	85-89, 341-345
CLO III-22	425	2	46-50, 344-348
CLO III-27	128	1	65-69
CLO III-52	446	1	113-117
CLO III-53	152	Not found	No location
CLO III-55	432	1	36-40
CLO III-57	154	Not found	No location
CLO III-58	150	Not found	No location
CLO III-63	154	Not found	No location
CLO III-74	153	1	53-57
CLO III-78	450	2	304-308, 335-339
CLO III-80	451	1	385-389

Table 5: The number and the position of the CCCTC motifs in selected CLO inserts.
* CLO II-7, and II-17 have not been sequenced yet.

Chapter 4: Discussion

4.1 Experimental Discussion

Boundary elements or insulators are an additional class of chromosomal *cis*-acting DNA regulatory elements capable of modulating gene expression, with two experimentally defined properties. First, an insulator may act as an enhancer-blocking element when it is positioned between an enhancer and its target promoter, without affecting the enhancer activity since this enhancer is still able to activate a non-insulated promoter (Bell *et al* 2001; Capelson and Corces 2004). Second, insulators are able to work as barriers protecting a transgene from chromosomal position effect when flanking it, allowing for position-independent gene expression. Almost all insulators possess both enhancer-blocking and barrier properties except those boundary elements identified in yeast which act primarily as barriers (Caperson and Corces 2004). Even though the two insulator properties are separated since each one depends on specific sequences for function, in some cases they seem physically linked since both shield the gene from the surrounding signals. The activity of these elements requires both *cis*- and *trans*- acting elements working together; different *cis* insulator DNA sequences bind different *trans* facilitating proteins such as CTCF protein, an evolutionarily conserved protein that is involved in the enhancer-blocking activity of most described vertebrate insulators. Plant cells resemble animal cells in the way they function and in the mechanisms of action involved in gene regulation overall; the similarity can be extended to the ability of some insulators from other species to function in *Arabidopsis* (Gudynaite-Savitch *et al*. 2009) this suggesting that plant may use insulators to regulate their gene expression as other

organisms do. Therefore, the hypothesis that plants must have *cis*-acting DNA sequences that interact with *trans*-acting sequences with insulator activity has been proposed.

The 19 putative insulator sequences selected were cloned into pB31 and tested for insulator activity by histochemical GUS staining. Each clone had a number of transgenic plants including insert CLO I-3, which had 49 transgenic plants (as shown in the Appendix **Table A1**). Studies have showed that the expression of the *GUS* gene is affected by various biochemical, molecular, and biological factors (Basu *et al.* 2004). For example, the expression of the *GUS* gene in different tissues at different developmental stages may be affected by the type of promoter used to derive its expression (Basu *et al.* 2004). Therefore, the stable expression of the *GUS* gene under the control of the CaMV*35S* promoter was detected in different tissues at different times. GUS staining results for different clones in pB31 vectors (**Table 3**) illustrated that, in some clones the *GUS* gene activity was observed in different plant tissues (flower, leaf, silique), which indicates that these clones have no insulator activity. Results for the clones that showed little or no staining indicated that these inserts might possess insulator activity. So far, results of the GUS staining have only identified one promising sequence (CLO I-3) that might possess insulator activity. Insert CLO I-3 had only two transgenic leaf samples stained blue out of 49 samples obtained from two different replicates. Each replicate had only two leaf samples such as leaf sample #19 and #45 that have only1 of the 3 leaf tissues in the same tube with a light GUS staining and even then only at the lowest level of detection ("1"). These had absolutely no GUS staining in the other transgenict issues tested (flowers and silique). The appearance of the blue dye in only two transgenic leaf samples may be due to chromosomal position effects as previously observed by Cai and

Shen (2001) who showed that insulators in transgenic constructs could not completely block misexpression resulting from inappropriate enhancer-promoter interactions in all transformants. In that study, the effect of *Drosophila* Su(Hw) insulator copy number on its insulator strength in *Drosophila* embryos was tested by inserting this insulator upstream of the *zerknult* enhancer VRE (ventral repression element) which mediates dorsal activation of the divergently transcribed *mini white* gene. Results obtained from their study showed that flanking the enhancer by one copy of the Su(Hw) insulator blocked misexpression up to 76% while flanking the enhancer by two copies of the Su(Hw) eliminated misexpression up to 87.5%, which implies that misexpression in transgenic constructs likely occurs as a result of chromosomal position effects and insulators cannot totally prevent inappropriate enhancer-promoter interactions in their transgenic system.

In addition, results from the GUS staining also identified two other sequences (CLO II-10 and CLO III-78) that might possess insulator activity. However even though both CLO II-10 and CLO III-78 had moderate sample numbers, the GUS staining results obtained from both of these two sequences are still ambiguous. For CLO II-10, plants from the first transformation had only two transgenic flower samples with blue dye, whereas most plant tissues from the second transformation stained blue. Similar observations were made with CLO III-78 transformants. PCR amplification for specific sequences from the samples of CLO II-10 and CLO III-78 transgenics that did not stain for GUS activity demonstrated that the desired sequences (GUS gene, 35S core promoter, CLO insert) were present. Therefore, the most likely scenario for this observation is that an error such as mislabeling of the *A. tumefaciens* culture or collected transgenic seeds

had occurred before they are screened for the *GUS* gene activity. Further transformations will be needed to resolve the contradictory results observed with CLO II-10 and CLO III-78.

GUS staining results identified several possible tissue-specific insulators such as CLO II-12, and CLO III-52 (**Table 3**) with no staining in siliques & seeds. One possible explanation for this phenomenon is that, these sequences probably might possess insulator activity but in specific tissues the level of some components such as binding proteins that interact with these sequences necessary for insulator activity were too low or absent. Thus, this prevents these sequences from working as insulator in these tissues. In order to confirm these results more plants arising from independent transformations will be needed.

For the promising candidate CLO I-3 DNA samples from second transformation were extracted from transgenic plant leaves for PCR amplification to first assess the success of DNA extraction by using SALK primers to amplify the *A. thaliana* housekeeping gene (SALK), and then to confirm the success of transformation by testing for the presence of the*35S* core promoter and the *GUS* gene. Confirming the presence of all of these sequences is considered important to define the insert's activity. The PCR results verify the presence of all desired sequences including: the insert CLO I-3. The same PCR testing method was performed on DNA samples that were extracted from transgenic plant leaves obtained from the first transformation event each of the other two sequences (CLO II-10 and CLOIII-78) that may possess insulator activity. Results from PCRs for both inserts confirmed the presence of sequences of the expected size.

Positive sequencing results confirmed that CLO I-3 was present in the extracted DNA samples of transgenic leaves (5/5 samples cloned), implying that it was successfully transformed. The other two fragments (II-10 and III-78) were not sent for sequencing since more assessment is needed to classify their activity.

For the GUS staining results of the 19 sequences sub-cloned into pB31, the question can be asked: why did the majority of sequences that were identified in the first selection system pC1 not work here in pB31? A possible answer to this question is that mutations in the 35S46 promoter or the *codA* gene might have happened during transformation or plant regeneration which allowed the plants to grow in the selective media even if these sequences had not possess insulator activity.

When CLO inserts were sequenced, many were of the expected size of ~150bp but others had sizes between 427- 451bp. We proposed that this class arose during ligation into pC1 (**Figure 14)**. Explaining this phenomenon required identifying all the external and internal *Eco*RI and *Bam*HI sites (**Figure 13**). Restriction mapping results demonstrated that most fragments contained the sites Eco-Bam-Eco-Bam (**Figure 13**), as expected for the ligation of multiple 150bp fragments. Multiple sequence alignments did reveal the PCR primer sequences associated with the restriction sites, as expected (**Table 2**). Results from the alignments of each long sequence with itself and with all others did not detect a common insert, suggesting that each long sequence consisted of three unique fragments assembled together during the ligation process. Three exceptions were observed: CLO II-12, CLO III-55 and CLO III-78 (**Figure 13)**. These have extra restriction sites for BamH1 and EcoR1 that may have arisen during the synthesis of the CLO insertions. In addition, observation of the different sizes for the long oligonucleotide

sequences raised this question: why all these sequences having different sizes ranging between 427-451bp, despite the fact that all of them expected to be 450bp long? A possible answer to this question is that these sequences might have resulted from the deletion of additional of DNA during PCR amplification that happened during library preparation.

To date all studies that have been performed on the fruit fly (*D. melanogaster*), vertebrates, and humans support the model that a *cis*-acting DNA insulator binds to a transacting protein e.g., CCCTC binding CTCF, which is the most common protein identified as involved in insulator activity. In vertebrates, this protein binds to three regularly spaced repeats of the core sequence CCCTC and thus was named CCCTC binding factor (Lobanenkov *et al.* 1990). All CLO sequences were scanned for CCCTC. The majority of them have at least one copy of this DNA motif (**Table 5)** including CLO I-3, the only sequence identified from the oligonucleotide library with possible enhancer-blocking activity. Thus CLO I-3 must possess a unique sequence that works instead of the CCCTC or a second *cis*-acting DNA that works with CCCTC to confer its insulator activity. Two sequences GCCCATACAT and ATCCATACAT, within UASrpg of *A. gossypii* (Steiner *et al.* 1994) previously identified as potential insulators, were not found within CLO I-3.Thus, more analysis should be carried out to identify the potential sequence in insert CLO I-3 since no homology with any known insulator sequence has been observed yet. An experimental approach may be necessary to identify the active motif(s), such as deletion analysis.

4.2 Future work

To be useful as a general insulator it must be demonstrated that CLO 1-3 and any other potential insulator sequences can function with different promoters and in different plant species. Thus the first step would be their cloning from pB31 into the vector pL1. As described in Section 3.1 for pL1, the napin seed-specific promoter replaces the CaMV*35S*46 promoter found in pB31 as the promoter driving GUS expression. Otherwise, the experimental approach is the same and if the fragments do indeed have insulator function, *GUS* will be expressed only in transgenic plant seeds but if do not possess insulator activity the expression of GUS gene will be appear in all plant tissues.

Furthermore, if the insert CLO I-3 worked as an insulator in these three selection systems it would be interesting to identify its potential insulator sequence. This could be possibly achieved by dividing the long sequence of the insert CLO I-3 into three sequences using *Eco*RI and *Bam*HI restriction enzymes to restrict this sequence at these specific sites, then cloning each sequence individually into a selection system such as pB31. Detecting the absent of activity of *GUS* gene in plant tissues will demonstrate which sequence among these three is conferring the insulation activity. If this approach did not work, another one could be performed using deletion analysis to detect the potential sequence in CLO I-3. As shown in **Table 5** CLO I-3 has two copies of the CCCTC sequence at 58-62 and 336-340bp thus upstream and downstream sequence deletions would be designed to identify the active sequence. The first deletion would be a downstream deletion since the first sequence could start from 1-62. The second deletion would be an upstream deletion since the second sequence will start from 63-340, and the third deletion would also be an upstream deletion since the sequence will start from 58-340. Cloning these sequences

separately into a selection vector might help to identify this sequence. In addition, cloning these fragments in both orientations would be valuable to test if insulators are able to work in both directions and to display the effect of the orientation on insulator activity. Testing oligonucleotide fragments in three different promoter/enhancer combinations in a model system will help to identify a true insulator that might have the potential to work within various genomic contexts, thus making it more appropriate for use in engineering of future transgenic lines. In addition, the same overall approach could be used to identify insulators in an *A. thaliana* genomic library.

Insulators from different species, such as UASrpg from *A. gossypii* and DEAD-1C from the human *T-cell receptor* α/δ *locus*, have the ability to work as enhancer-blocking elements in *A. thaliana* (Gudynaite-Savitch *et al.* 2009; Singer *et al* 2011), proving that the insulator activity is conserved among species. Thus the insulator(s) that we have identified may function in various plants such as tobacco (Wilkinson *et al*. 1996) and maize where the CaMV*35S* promoter has been proven to work. This would be especially desirable for transgenic research. So further work may involve transformation and efficacy assays in species other than *A. thaliana* that have economic importance, such as tobacco, tomato and soybean, by using *A. tumefaciens* (Mayo *et al* 2006; McCormick *et al*1991) or biolistic methods. This approach could be extended to yeast and fungi as well.

Significant DNA sequence similarity has been observed among insulators identified from various organisms (Bell *et al*. 2001), which might not support the notion of the existence of a universal insulator; thereforethere may be no single universal insulator.

Several insulators are known to interact with specific proteins. Therefore, the long term goal here is an understanding of the mechanism underling insulator activity at the

molecular level since this will contribute to the fundamental understanding of spatial and temporal regulation of gene expression in plants. This goal could be reached first by identifying an efficient insulator that could function in *Arabidopsis* as well as with other plant species of agricultural importance. Second, use selected insulators to test insulator proposed models (barrier and enhancer-blocking) by characterizing their roles in the regulation of specific genes in *A. thaliana*. Third, isolate the insulator-binding proteins and their associated proteins mediating the function of insulator by using selected insulators as probes. Standard protein interaction techniques such as co-immunoprecipitation and yeast two-hybrid assays may be used to identify the protein cofactors necessary for conferring insulator activity to the identified insulator sequences. Also, the formation of the insulator-nuclear protein complex could be tested by an electrophoretic mobility shift assay (EMSA). The ability of DNA sequence(s) with insulator activity to interact with nuclear proteins and/or transcriptional factors would be tested in a yeast one-hybrid screen in order to identify *A. thaliana* genes encoding specific protein(s) involved in direct interaction with insulator(s).

4.3 Conclusion

Identifying sequence(s) possessing insulator activity in a model system such as *A. thaliana* will potentially have a huge impact not only on the improvement of plant biotechnological research of important traits in crops but also any biological study or industry requiring precise expression patterns of transformed genes. Insulators with their ability to protect transgenes from chromosomal position effects will be utilized to increase the efficiency of genetic engineering resulting in transgenic organisms with

more stable, predictable traits. In addition, by blocking inappropriate enhancer-promoter interactions, insulator could be used in the production of biopharmaceuticals, edible vaccines, and antibodies in targeted tissues of some crops and legume species (Gudynaite-Savitch *et al.* 2009). Moreover, they could help in the process of enhancing plant transformation vectors and to increase the performance of several traits at once by designing transformation vectors that harbor multiple transcriptional gene units (Singer *et al.* 2011). This study is an early step in a longer process of validating the activity of these sequences in other promoter/reporter systems and other model organisms, identifying associated protein cofactors, their mode of action, and ultimately fully characterizing and using a functional plant insulator sequence for transgenic engineering.

Chapter 5: References

Allen, G.C., Spiker, S., Thompson, W.F., 2000.Use of matrix attachment regions (MARs) to minimize transgene silencing. Plant Molecular Biology 43, 361–376.

Anthony, A., Blaxter, M., 2007.Association of the matrix attachment region recognition signature with coding regions in *Caenorhabditis elegans*. BioMedCenter Genomics 8, 418.

Arumugam, P.I., Urbinati, F., Velu, C.S., Higashimoto, T., Grimes, H.L., Malik, P., 2009. The 3′ Region of the Chicken Hypersensitive Site-4 Insulator Has Properties Similar to Its Core and Is Required for Full Insulator Activity. Public Library of Science ONE 4, e6995.

Bao, L., Zhou, M., Cui, Y., 2008. CTCFBSDB: a CTCF-binding site database for characterization of vertebrate genomic insulators. Nucleic Acids Research 36, D83–D87.

Barges, S., Mihaly, J., Galloni, M., Hagstrom, K., Müller, M., Shanower, G., Schedl, P., Gyurkovics, H., Karch, F., 2000. The Fab-8 boundary defines the distal limit of the bithorax complex iab-7 domain and insulates iab-7 from initiation elements and a PRE in the adjacent iab-8 domain. Development 127, 779–790.

Basu, C., Kausch, A.P., Chandlee, J.M., 2004. Use of β-glucuronidase reporter gene for gene expression analysis in turf grasses. Biochemical and Biophysical Research Communications 320, 7–10.

Baxley, R.M., Soshnev, A.A., Koryakov, D.E., Zhimulev, I.F., Geyer, P.K., 2011. The role of the Suppressor of Hairy-wing insulator protein in Drosophila oogenesis.Development Biology 356, 398–410.

Bell, A.C., Felsenfeld, G., 1999. Stopped at the border: boundaries and insulators. Current Opinion in Genetics & Development 9, 191–198.

Bell, A.C., West, A.G., Felsenfeld, G., 1999. The Protein CTCF Is Required for the Enhancer Blocking Activity of Vertebrate Insulators. Cell 98, 387–396.

Bell, A.C., West, A.G., Felsenfeld, G., 2001. Insulators and Boundaries: Versatile Regulatory Elements in the Eukaryotic Genome. Science 291, 447–450.

Belozerov, V.E., 2003. A novel boundary element may facilitate independent gene regulation in the Antennapedia complex of Drosophila. The EMBO Journal 22, 3113–3121.

Benfey, P.N., Ren, L., Chua, N.H., 1990.Tissue-specific expression from CaMV 35S enhancer subdomains in early stages of plant development. The EMBO Journal 9, 1677–1684.

Bi, X., Broach, J.R., 1999.UASrpg Can Function as a Heterochromatin Boundary Element in Yeast. Genes Development 13, 1089–1101.

Bi, X., Broach, J.R., 2001.Chromosomal boundaries in *S. cerevisiae*. Current Opinion in Genetics & Development 11, 199–204.

Brasset, E., Vaury, C., 2005. Insulators are fundamental components of the eukaryotic genomes. Heredity 94, 571–576.

Bushey, A.M., Dorman, E.R., Corces, V.G., 2008. Chromatin insulators: regulatory mechanisms and epigenetic inheritance. Molecular Cell 32, 1–9.

Byrd, K., Corces, V.G., 2003. Visualization of Chromatin Domains Created by the Gypsy Insulator of Drosophila. The Journal of Cell Biology 162, 565–574.

Cai, H.N., Shen, P., 2001. Effects of Cis Arrangement of Chromatin Insulators on Enhancer-Blocking Activity. Science 291, 493–495.

Capelson, M., Corces, V.G., 2004.Boundary elements and nuclear organization. Biology of the Cell 96, 617–629.

Capelson, M., Corces, V.G., 2005. The ubiquitin ligase dTopors directs the nuclear organization of a chromatin insulator. Molecular Cell20, 105–116.

Capelson, M., Corces, V.G., 2004.Boundary elements and nuclear organization. Biology of the Cell 96, 617–629.

Cavalieri, V., Melfi, R., Spinelli, G., 2009. Promoter activity of the sea urchin (*Paracentrotus lividus*) nucleosomal H3 and H2A and linker H1 α-histone genes is modulated by enhancer and chromatin insulator. Nucleic Acids Research 37, 7407–7415.

Celniker, S.E., Drewell, R.A., 2007. Chromatin looping mediates boundary element promoter interactions. BioEssays 29, 7–10.

Chang, H.Y., Cuvier, O., Dekker, J., 2009. Gene dates, parties and galas. Symposium on Chromatin Dynamics and Higher Order Organization 10, 689–693.

Chen, S., Corces, V.G., 2001.The *Gypsy* Insulator of Drosophila Affects Chromatin Structure in a Directional Manner. Genetics 159, 1649–1658.

Ciavatta, D., Kalantry, S., Magnuson, T., Smithies, O., 2006. A DNA Insulator Prevents Repression of a Targeted X-Linked Transgene but Not Its Random or Imprinted X Inactivation. Proceedings of the National Academy of Science 103, 9958–9963.

Clark, A.J., Bissinger, P., Bullock, D.W., Damak, S., Wallace, R., Whitelaw, C.B., Yull, F., 1994.Chromosomal position effects and the modulation of transgene expression. Reproduction, Fertility and Development 6, 589–598.

Conte, C., Dastugue, B., Vaury, C., 2002.Coupling of Enhancer and Insulator Properties Identified in Two Retrotransposons Modulates Their Mutagenic Impact on Nearby Genes. Molecular and Cellular Biology 22, 1767–1777.

Defossez, P. A., Gilson, E., 2002. The Vertebrate Protein CTCF Functions as an Insulator in *Saccharomyces cerevisiae*. Nucleic Acids Research 30, 5136–5141.

Donze, D., Adams, C.R., Rine, J., Kamakaka, R.T., 1999. The boundaries of the silenced HMR domain in *Saccharomyces cerevisiae*. Genes & Development 13, 698–708.

Du, T. T., Huang, Q.-H., 2007. The roles of histone lysine methylation in epigenetic regulation. Yi Chuan 29, 387–392.

Eriksson, P., Alipour, H., Adler, L., Blomberg, A., 2000. Rap1p-binding sites in the S*saccharomyces cerevisiae* GPD1 promoter are involved in its response to NaCl. The Journal of Biological Chemistry 275, 29368–29376.

Fourel, G., Boscheron, C., Revardel, manuelle, Lebrun, lonore, Hu, Y.-F., Simmen, K.C., Mller, K., Li, R., Mermod, N., Gilson, |ric, 2001. An activation-independent role of transcription factors in insulator function. EMBO Reports 2, 124–132.

Fourel, G., Magdinier, F., Gilson, É., 2004. Insulator dynamics and the setting of chromatin domains. BioEssays 26, 523–532.

Fourel, G., Revardel, E., Koering, C.E., Gilson, E., 1999. Cohabitation of insulators and silencing elements in yeast subtelomeric regions.The EMBO Journal 18, 2522–2537.

G|ndr, A., Ohlsson, R., 2008.Chromatin insulators and cohesins. The EMBO reports 9, 327–329.

Galliano, H., Müller, A.E., Lucht, J.M., Meyer, P., 1995. The transformation booster sequence from *Petunia hybrida* is a retrotransposon derivative that binds to the nuclear scaffold. Molecular Genetics and Genomic 247, 614–622.

Gaszner, M., Felsenfeld, G., 2006.Insulators: exploiting transcriptional and epigenetic mechanisms. Nature Reviews Genetics 7, 703–713.

Gerasimova, T.I., Byrd, K., Corces, V.G., 2000. A Chromatin Insulator Determines the Nuclear Localization of DNA. Molecular Cell 6, 1025–1035.

Geyer, P.K., 1997. The role of insulator elements in defining domains of gene expression. Current Opinion in Genetics & Development 7, 242–248.

Ghosh, D., Gerasimova, T.I., Corces, V.G., 2001. Interactions between the Su(Hw) and Mod(mdg4) proteins required for gypsy insulator function. The EMBO Journal 20, 2518–2527.

Golovnin, A., Melnikova, L., Volkov, I., Kostuchenko, M., Galkin, A.V., Georgiev, P., 2008. Insulator bodies are aggregates of proteins but not of insulators. The EMBO Reports 9, 440–445.

Gombert, W.M., Farris, S.D., Rubio, E.D., Morey-Rosler, K.M., Schubach, W.H., Krumm, A., 2003. The c-myc insulator element and matrix attachment regions define the c-myc chromosomal domain. Molecular and Cellular Biology 23, 9338–9348.

Gudynaite-Savitch, L., Johnson, D.A., Miki, B.L.A., 2009.Strategies to mitigate transgene–promoter interactions. Plant Biotechnology Journal 7, 472–485.

Guo, C., Yoon, H.S., Franklin, A., Jain, S., Ebert, A., Cheng, H. L., Hansen, E., Despo, O., Bossen, C., Vettermann, C., Bates, J.G., Richards, N., Myers, D., Patel, H., Gallagher, M., Schlissel, M.S., Murre, C., Busslinger, M., Giallourakis, C.C., Alt, F.W., 2011. CTCF-binding elements mediate control of V(D)J recombination. Nature 477, 424–430.

Hily, J. M., Singer, S., Yang, Y., Liu, Z., 2009. A transformation booster sequence from Petunia hybrid functions as an enhancer-blocking insulator in *Arabidopsis thaliana*. Plant Cell Reports 28, 1095–1104.

Holohan, E.E., Kwong, C., Adryan, B., Bartkuhn, M., Herold, M., Renkawitz, R., Russell, S., White, R., 2007. CTCF Genomic Binding Sites in Drosophila and the Organisation of the Bithorax Complex. Public Library of Science Genet 3.

Idrissi, F.Z., Fernández-Larrea, J.B., Piña, B., 1998. Structural and functional heterogeneity of Rap1p complexes with telomeric and UASrpg-like DNA sequences. The Journal of Molecular Biology 284, 925–935.

Jeong, S., Pfeifer, K., others, 2004.Shifting insulator boundaries. Nature Genetics 36, 1036–1037.

Kadauke, S., Blobel, G.A., 2009. Chromatin loops in gene regulation. Biochimica et BiophysicaActa (BBA) - Gene Regulatory Mechanisms 1789, 17–25.

Karcher, S.J., 2002. Blue plants: transgenic plants with the GUS reporter gene, in: Tested Studies for Laboratory Teaching. Proceedings of the 23rd Workshop., MA O'Donnell, Ed (Conference of the Association for Biology Laboratory Education (ABLE) pp 29–42.

Kellum, R., Schedl, P., 1991. A position-effect assay for boundaries of higher order chromosomal domains. Cell 64, 941–950.

Kellum, R., Schedl, P., 1992. A Group of *scs* Elements Function as Domain Boundaries in an Enhancer-Blocking Assay.Molecular and Cellular Biology 12, 2424–2431.

Kim, T.H., Abdullaev, Z.K., Smith, A.D., Ching, K.A., Loukinov, D.I., Green, R.D., Zhang, M.Q., Lobanenkov, V.V., Ren, B., 2007.Analysis of the Vertebrate Insulator Protein CTCF-Binding Sites in the Human Genome. Cell 128, 1231–1245.

Kouzarides, T., 2007.Chromatin Modifications and Their Function. Cell 128, 693–705.

Kravchenko, E., Savitskaya, E., Kravchuk, O., Parshikov, A., Georgiev, P., Savitsky, M., 2005. Pairing Between Gypsy Insulators Facilitates the Enhancer Action in Trans Throughout the Drosophila Genome. Molecular and Cellular Biology 25, 9283–9291.

Kyrchanova, O.V., Ivleva, T.A., Georgiev, P.G., 2011. Interacting insulators from the Drosophila melanogaster bithorax complex can form independent expression domains. Genetika 47, 1586–1595.

Labrador, M., Corces, V.G., 2002.Setting the Boundaries of Chromatin Domains and Nuclear Organization. Cell 111, 151–154.

Law, J.A., Jacobsen, S.E., 2010. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nature Reviews Genetics 11, 204–220.

Li, M., Belozerov, V.E., Cai, H.N., 2008.Analysis of chromatin boundary activity in Drosophila cells. BMC Molecular Biology 9, 109.

Ling, J., Ainol, L., Zhang, L., Yu, X., Pi, W., Tuan, D., 2004. HS2 Enhancer Function Is Blocked by a Transcriptional Terminator Inserted Between the Enhancer and the Promoter. The Journal of Biological Chemistry 279, 51704–51713.

Litt, M.D., 2001. Transitions in histone acetylation reveal boundaries of three separately regulated neighboring loci. The EMBO Journal 20, 2224–2235.

Litt, M.D., Simpson, M., Gaszner, M., Allis, C.D., Felsenfeld, G., 2001. Correlation between histone lysine methylation and developmental changes at the chicken beta-globin locus. Science 293, 2453–2455.

Lobanenkov, V.V., Nicolas, R.H., Adler, V.V., Paterson, H., Klenova, E.M., Polotskaja, A.V., Goodwin, G.H., 1990. A novel sequence-specific DNA binding protein which interacts with three regularly spaced direct repeats of the CCCTC-motif in the 5'-flanking sequence of the chicken c-myc gene. Oncogene 5, 1743–1753.

Lourdes Valenzuela, Rohinton, T.K., 2012. Chromatin Insulators. Docstoc.com. URL http://www.docstoc.com/docs/103974357/Chromatin-Insulators.

Lynch, P.J., Rusche, L.N., 2009. A Silencer Promotes the Assembly of Silenced Chromatin Independently of Recruitment. Molecular and Cellular Biology 29, 43–56.

Mahmoudi, T., Katsani, K.R., Verrijzer, C.P., 2002. GAGA can mediate enhancer function *in trans* by linking two separate DNA molecules. The EMBO Journal 21, 1775–1781.

Majumder, P., Cai, H.N., 2003. The Functional Analysis of Insulator Interactions in the Drosophila Embryo.Proceeding of National Academy of Science 100, 5223–5228.

Majumder, P., Roy, S., Belozerov, V.E., Bosu, D., Puppali, M., Cai, H.N., 2009. Diverse Transcription Influences Can Be Insulated by the Drosophila SF1 Chromatin Boundary. Nucleic Acids Research 37, 4227–4233.

Markstein, M., Pitsouli, C., Villalta, C., Celniker, S.E., Perrimon, N., 2008. Exploiting position effects and the gypsy retrovirus insulator to engineer precisely expressed transgenes. Nature Genetics 40, 476–483.

Mayo, K.J., Gonzales, B.J., Mason, H.S., 2006. Genetic transformation of tobacco NT1 cells with Agrobacterium tumefaciens. Natural Protocols 1, 1105–11.

McCormick, S. 1991.Transformation of tomato with *Agrobacterium tumefaciens.* Plant Tissue Culture Manual 1–9.

Melnikova, L., Juge, F., Gruzdeva, N., Mazur, A., Cavalli, G., Georgiev, P., 2004. Interaction Between the GAGA Factor and Mod(mdg4) Proteins Promotes Insulator Bypass in Drosophila. Proceedings of National Academy of Science 101, 14806–14811.

Mitsuhara, I., Ugaki, M., Hirochika, H., Ohshima, M., Murakami, T., Gotoh, Y., Katayose, Y., Nakamura, S., Honkura, R., Nishimiya, S., Ueno, K., Mochizuki, A., Tanimoto, H., Tsugawa, H., Otsuki, Y., Ohashi, Y., 1996. Efficient promoter cassettes for enhanced expression of foreign genes in dicotyledonous and monocotyledonous plants. Plant Cell Physiol. 37, 49–59.

Mongelard, F., Corces, V.G., 2001. Two insulators are not better than one. Nature Structural Biology 8, 192–194.

Muravyova, E., Golovnin, A., Gracheva, E., Parshikov, A., Belenkaya, T., Pirrotta, V., Georgiev, P., 2001. Loss of Insulator Activity by Paired Su(Hw) Chromatin Insulators. Science 291, 495–498.

Ohlsson, R., Lobanenkov, V., Klenova, E., 2010. Does CTCF mediate between nuclear organization and gene expression? BioEssays 32, 37–50.

Ong, C.T., Corces, V.G., 2009. Insulators as mediators of intra-and inter-chromosomal interactions: a common evolutionary theme. The Journal of Biology 8, 73.

Pai, C. Y., Lei, E.P., Ghosh, D., Corces, V.G., 2004. The Centrosomal Protein CP190 Is a Component of the gypsy Chromatin Insulator. Molecular Cell 16, 737–748.

Petrascheck, M., Escher, D., Mahmoudi, T., Verrijzer, C.P., Schaffner, W., Barberis, A., 2005.DNA Looping Induced by a Transcriptional Enhancer in Vivo. Nucleic Acids Research 33, 3743–3750.

Pfluger, J., Wagner, D., 2007. Histone modifications and dynamic regulation of genome accessibility in plants. Current Opinion in Plant Biology 10, 645–652.

Pikaart, M.J., Recillas-Targa, F., Felsenfeld, G., 1998. Loss of transcriptional activity of a transgene is accompanied by DNA methylation and histone deacetylation and is prevented by insulators. Genes Development 12, 2852–2862.

Ramezani, A., Hawley, T.S., Hawley, R.G., 2008. Combinatorial incorporation of enhancer-blocking components of the chicken beta-globin 5'HS4 and human T-cell receptor alpha/delta BEAD-1 insulators in self-inactivating retroviral vectors reduces their genotoxic potential. Stem Cells 26, 3257–3266.

Ramos, E., Ghosh, D., Baxter, E., Corces, V.G., 2006. Genomic Organization of Gypsy Chromatin Insulators in *Drosophila melanogaster*. Genetics 172, 2337–2349.

Recillas-Targa, F., 2002. Position-effect protection and enhancer blocking by the chicken beta -globin insulator are separable activities. Proceedings of the National Academy of Sciences 99, 6883–6888.

Rivella, S., Callegari, J.A., May, C., Tan, C.W., Sadelain, M., 2000. The cHS4 insulator increases the probability of retroviral expression at random chromosomal integration sites. Journal of Virology 74, 4679–4687.

Salinas, J., Sánchez-Serrano, J.J., 2006. Arabidopsis Protocols. Humana Press.

Savitskaya, E., Melnikova, L., Kostuchenko, M., Kravchenko, E., Pomerantseva, E., Boikova, T., Chetverina, D., Parshikov, A., Zobacheva, P., Gracheva, E., Galkin, A., Georgiev, P., 2006. Study of long-distance functional interactions between Su(Hw) insulators that can regulate enhancer-promoter communication in Drosophila melanogaster. Molecular and Cellular Biology 26, 754–761.

Scott, K.C., Taubman, A.D., Geyer, P.K., 1999. Enhancer Blocking by the Drosophila *Gypsy* Insulator Depends Upon Insulator Anatomy and Enhancer Strength. Genetics 153, 787–798.

Sieburth, L.E., Meyerowitz, E.M., 1997. Molecular dissection of the AGAMOUS control region shows that cis elements for spatial regulation are located intragenically. The Plant Cell 9, 355–65.

Singer, S.D. and Cox, K.D. 2012. A *gypsy*-like sequence from *Arabidopsis thaliana* exhibits enhancer-blocking activity in transgenic plants. Journal of Plant Biochemistry and Biotechnology DOI 10.1007/s13562-012-0108-3.

Singer, S.D., Cox, K.D., Liu, Z., 2010. Both the constitutive cauliflower mosaic virus 35S and tissue-specific AGAMOUS enhancers activate transcription autonomously in Arabidopsis thaliana. Plant Molecular Biology 74, 293–305.

Singer, S.D., Cox, K.D., Liu, Z., 2011. Enhancer-promoter interference and its prevention in transgenic plants. The Plant Cell Reporter 30, 723–731.

Singer, S.D., Hily, J. M., Liu, Z., 2009. A 1-kb Bacteriophage Lambda Fragment Functions as an Insulator to Effectively Block Enhancer–Promoter Interactions in Arabidopsis thaliana. The Plant Molecular Biology Reporter 28, 69–76.

Steiner, S., Philippsen, P., 1994.Sequence and promoter analysis of the highly expressed TEF gene of the filamentous fungus *Ashbya gossypii*.Molecular and General Genetics. 242, 263–271.

Steinwaerder, D.S., Lieber, A., 2000. Insulation from viral transcriptional regulatory elements improves inducible transgene expression from adenovirus vectors *in vitro* and *in vivo*. Gene Therapy 7, 556–567.

Szostková, M., Horáková, D., 1998.The effect of plasmid DNA sizes and other factors on electrotransformation of Escherichia coli JM109.Bioelectrochemistry and Bioenergetics 47, 319–323.

van der Geest, A.H., Hall, T.C., 1997. The beta-phaseolin 5' matrix attachment region acts as an enhancer facilitator. Plant Molecular Biology 33, 553–557.

van der Vlag, J., den Blaauwen, J.L., Sewalt, R.G., van Driel, R., Otte, A.P., 2000. Transcriptional repression mediated by *polycomb* group proteins and other chromatin-associated repressors is selectively blocked by insulators. The Journal of Biological Chemistry 275, 697–704.

Wallace, J.A., Felsenfeld, G., 2007. We gather together: insulators and genome organization. Curr.Opin. Genetic Development 17, 400–407.

Walters, M.C., Fiering, S., Bouhassira, E.E., Scalzo, D., Goeke, S., Magis, W., Garrick, D., Whitelaw, E., Martin, D.I.K., 1999.The Chicken B-Globin 5′HS4 Boundary Element Blocks Enhancer-Mediated Suppression of Silencing. Molecular and Cellular Biology 19, 3714–3726.

Wei, G.H., Liu, D.P., Liang, C.C., 2005. Chromatin domain boundaries: insulators and beyond. Cell Research 15, 292–300.

West, A.G., Gaszner, M., Felsenfeld, G., 2002. Insulators: many functions, many mechanisms. Genes Development 16, 271–288.

Wilkinson, J.E., Twell, D., Lindsey, K., 1997. Activities of CaMV 35S and Nos Promoters in Pollen: Implications for Field Release of Transgenic Plants. The Journal of Experimental Botany 48, 265–275.

Xin, L., Liu, D. P., Ling, C.-C., 2003.A hypothesis for chromatin domain opening. BioEssays 25, 507–514.

Yang, Y., Singer, S.D., Liu, Z., 2010.Evaluation and comparison of the insulation efficiency of three enhancer-blocking insulators in plants. Plant Cell, Tissue and Organ Culture 105, 405–414.

Yao, S., Osborne, C.S., Bharadwaj, R.R., Pasceri, P., Sukonnik, T., Pannell, D., Recillas-Targa, F., West, A.G., Ellis, J., 2003. Retrovirus Silencer Blocking by the cHS4 Insulator Is CTCF Independent. Nucleic Acids Research 31, 5317–5323.

Yusufzai, T.M., Felsenfeld, G., 2004. The 5′-HS4 Chicken B-Globin Insulator Is a CTCF-Dependent Nuclear Matrix-Associated Element. Proceeding of National Academy of Science 101, 8620–8624.

Yusufzai, T.M., Tagami, H., Nakatani, Y., Felsenfeld, G., 2004. CTCF tethers an insulator to subnuclear sites, suggesting shared insulator mechanisms across species. Molecular Cell 13, 291–298.

Zhong, X.P., Krangel, M.S., 1997. An enhancer-blocking element between alpha and delta gene segments within the human T cell receptor alpha/delta locus. Proceeding of National Academy of Science 94, 5219–5224.

Zlatanova, J., Caiafa, P., 2009. CTCF and Its Protein Partners: Divide and Rule? The journal of Cell Science 122, 1275–1284.

Appendix

Construct	Size	Date of GUS staining	Plant	Flower	Leaf 1	Leaf 2	Silique 1	Silique 2
CLO I-3 (pB31: pCAM 1300-35S46-GUS)	438	12-Mar	1	0	0	0	0	0
Total number of samples =49		12-Mar	2	0	0	0	0	0
		12-Mar	3	0	0	0	0	0
		12-Mar	4	0	0	0	0	0
		12-Mar	5	0	0	0	0	0
		12-Mar	6	0	0	0	0	0
		12-Mar	7	0	0	0	0	0
		12-Mar	8	0	0	0	0	0
		12-Mar	9	0	0	0	0	0
		12-Mar	10	0	0	0	0	0
		12-Mar	11	0	0	0	0	0
		12-Mar	12	0	0	0	0	0
		12-Mar	13	0	0	0	0	0
		12-Mar	14	0	0	0	0	0
		12-Mar	15	0	0	0	0	0
		12-Mar	16	0	0	0	0	0
		13-Mar	17	0	0	0	0	0
		13-Mar	18	0	0	0	0	0
		13-Mar	19	0	0	1	0	0
		13-Mar	20	0	0	0	0	0
		13-Mar	21	0	0	0	0	0
		13-Mar	22	0	0	0	0	0
		13-Mar	23	0	0	0	0	0
		13-Mar	24	0	0	0	0	0
		01-Oct	25	0	0	0	0	0
		01-Oct	26	0	0	0	0	0
		01-Oct	27	0	0	0	0	0
		01-Oct	28	0	0	0	0	0
		01-Oct	29	0	0	0	0	0
		01-Oct	30	0	0	0	0	0
		01-Oct	31	0	0	0	0	0
		01-Oct	32	0	0	0	0	0
		01-Oct	33	0	0	0	0	0
		01-Oct	34	0	0	0	0	0
		01-Oct	35	0	0	0	0	0
		01-Oct	36	0	0	0	0	0
		01-Oct	37	0	0	0	0	0
		01-Oct	38	0	0	0	0	0
		01-Oct	39	0	0	0	0	0
		01-Oct	40	0	0	0	0	0
		01-Oct	41	0	0	0	0	0
		01-Oct	42	0	0	0	0	0
		01-Oct	43	0	0	0	0	0
		01-Oct	44	0	0	0	0	0
		01-Oct	45	0	2	0	0	0
		01-Oct	46	0	0	0	0	0
		01-Oct	47	0	0	0	0	0
		01-Oct	48	0	0	0	0	0
		01-Oct	49	0	0	0	0	0
Total number of samples with GUS staining				0	1	1	0	0

CLO II-3 (pB31: pCAM 1300-35S46-GUS)	**151**	13-Mar	**1**	1	2	ND	2	ND
Total number of samples =14		13-Mar	**2**	1	0	ND	1	ND
		13-Mar	**3**	1	0	ND	0	ND
		nov,2010	**4**	0	0	0	0	0
		nov,2010	**5**	0	0	0	0	0
		nov,2010	**6**	0	3	1	3	3
		nov,2010	**7**	1	0	0	1	0
		nov,2010	**8**	0	2	1	1	2
		nov,2010	**9**	1	2	2	3	3
		nov,2010	**10**	1	0	0	0	0
		nov,2010	**11**	0	3	2	3	3
		nov,2010	**12**	0	0	0	0	0
		nov,2010	**13**	1	0	3	0	0
		nov,2010	**14**	0	0	1	0	ND
Total number of samples with GUS staining				**7**	**5**	**6**	**7**	**4**
CLO II-7 (pB31: pCAM 1300-35S46-GUS)	**~150**	Oct,2010	**1**	0	0	0	0	0
Total number of samples =30		Oct,2010	**2**	0	1	2	1	3
		Oct,2010	**3**	0	1	2	1	1
		Oct,2010	**4**	0	2	2	1	0
		Oct,2010	**5**	1	0	0	0	1
		Oct,2010	**6**	0	3	0	0	0
		Oct,2010	**7**	0	0	0	0	0
		Oct,2010	**8**	0	0	0	0	0
		Oct,2010	**9**	0	0	1	0	0
		Oct,2010	**10**	0	0	0	0	0
		Oct,2010	**11**	0	0	0	0	0
		Oct,2010	**12**	0	0	0	0	0
		Oct,2010	**13**	0	0	0	0	0
		Oct,2010	**14**	1	0	0	0	0
		Oct,2010	**15**	0	0	0	0	0
		Oct,2010	**16**	DNG	DNG	DNG	DNG	DNG
		Oct,2010	**17**	0	1	1	0	0
		19-Jul	**18**	1	2	2	0	1
		19-Jul	**19**	1	2	1	0	1
		19-Jul	**20**	1	0	0	0	1
		19-Jul	**21**	1	1	0	0	0
		19-Jul	**22**	1	0	0	0	0
		19-Jul	**23**	0	1	1	0	1
		19-Jul	**24**	1	0	1	0	0
		19-Jul	**25**	0	0	0	0	0
		19-Jul	**26**	1	0	0	0	0
		19-Jul	**27**	0	0	0	0	0
		19-Jul	**28**	1	0	1	0	0
		19-Jul	**29**	1	2	0	0	0
		19-Jul	**30**	0	1	0	0	0
Total number of samples with GUS staining				**11**	**11**	**10**	**3**	**7**

CLO II-10 (pB31: pCAM 1300-35S46-GUS)	**154**	23-Mar	**1**	0	0	ND	0	ND
Total number of samples =36		23-Mar	**2**	0	0	ND	0	ND
		23-Mar	**3**	DNG	DNG	DNG	DNG	DNG
		23-Mar	**4**	DNG	DNG	DNG	DNG	DNG
		23-Mar	**5**	1	0	ND	0	ND
		23-Mar	**6**	0	0	ND	0	ND
		23-Mar	**7**	DNG	DNG	DNG	DNG	DNG
		23-Mar	**8**	0	0	ND	0	ND
		07-Apr	**9**	1	0	0	0	0
		07-Apr	**10**	0	0	0	0	0
		07-Apr	**11**	0	0	0	0	0
		07-Apr	**12**	DNG	DNG	DNG	DNG	DNG
		07-Apr	**13**	0	0	0	0	0
		07-Apr	**14**	0	0	0	0	0
		07-Apr	**15**	0	0	0	0	0
		19-Jul	**16**	1	1	0	0	0
		19-Jul	**17**	1	1	0	0	0
		19-Jul	**18**	1	1	0	0	2
		19-Jul	**19**	1	0	1	0	0
		19-Jul	**20**	1	0	0	0	0
		19-Jul	**21**	1	2	0	0	3
		19-Jul	**22**	1	0	0	0	3
		19-Jul	**23**	1	1	0	0	3
		19-Jul	**24**	1	0	0	0	3
		19-Jul	**25**	1	1	0	0	3
		19-Jul	**26**	1	1	0	0	3
		19-Jul	**27**	1	0	0	0	3
		19-Jul	**28**	1	0	0	0	3
		19-Jul	**29**	1	0	0	0	0
		19-Jul	**30**	1	1	0	0	3
		19-Jul	**31**	1	0	0	0	3
		19-Jul	**32**	1	0	1	0	3
		19-Jul	**33**	1	1	0	0	0
		19-Jul	**34**	1	0	0	0	0
		19-Jul	**35**	1	1	0	0	0
		19-Jul	**36**	1	1	0	0	0
Total number of samples with GUS staining				**23**	**11**	**2**	**0**	**12**
CLO II-12 (pB31: pCAM 1300-35S46-GUS)	**427**	23-Mar	**1**	0	0	2	0	0
Total number of samples =15		23-Mar	**2**	0	0	3	0	0
		23-Mar	**3**	0	0	0	0	All dry
		23-Mar	**4**	0	0	0	0	0
		07-Apr	**5**	0	0	0	0	0
		07-Apr	**6**	0	0	1	0	0
		07-Apr	**7**	DNG	DNG	DNG	0	0
		07-Apr	**8**	DNG	DNG	DNG	0	0
		Mar30 ,2011	**9**	1	0	0	0	0
		Mar30 ,2011	**10**	1	0	0	0	0
		Mar30 ,2011	**11**	1	0	0	0	0
		Mar30 ,2011	**12**	1	0	0	0	0
		Mar30 ,2011	**13**	1	0	0	0	0
		2011.Apri	**14**	0	0	0	0	0
		2011.Apri	**15**	0	0	0	0	0
Total number of samples with GUS staining				**5**	**0**	**3**	**0**	**0**

CLO III-4 (pB31: pCAM 1300-35S46-GUS)	440	Nov,2010	1	1	1	0	0	0
Total number of samples =28		Nov,2010	2	1	1	0	0	0
		Nov,2010	3	1	0	0	0	0
		Nov,2010	4	1	0	0	0	0
		Nov,2010	5	1	0	0	0	0
		Nov,2010	6	1	0	0	0	0
		Nov,2010	7	1	0	0	ND	0
		Nov,2010	8	1	0	0	ND	0
		Nov,2010	9	1	0	0	0	0
		Nov,2010	10	1	0	0	ND	0
		Nov,2010	11	1	0	0	0	1
		Nov,2010	12	0	0	0	0	0
		Nov,2010	13	1	1	0	0	0
		Nov,2010	14	1	0	0	0	0
		Nov,2010	15	1	0	0	0	0
		Nov,2010	16	1	0	0	0	0
		Nov,2010	17	0	1	0	0	1
		Nov,2010	18	1	0	0	0	0
		Nov,2010	19	0	0	0	0	0
		Nov,2010	20	1	0	0	0	0
		Nov,2010	21	1	0	0	0	1
		Nov,2010	22	0	0	0	0	0
		Nov,2010	23	0	0	0	0	0
		Nov,2010	24	0	0	0	0	0
		Nov,2010	25	0	0	0	0	0
		Nov,2010	26	0	0	0	0	0
		Nov,2010	27	0	0	0	0	0
		Nov,2010	28	0	0	0	0	0
Total number of samples with GUS staining				18	4	0	0	3
CLO III-17 (pB31: pCAM 1300-35S46-GUS)	443	May,2010	1	1	0	1	0	0
Total number of samples =30		May,2010	2	0	0	1	0	0
		May,2010	3	0	0	0	0	0
		May,2010	4	1	1	0	0	0
		May,2010	5	1	0	0	0	0
		May,2010	6	1	0	1	0	0
		May,2010	7	1	1	0	0	0
		May,2010	8	1	1	0	0	1
		May,2010	9	1	0	1	0	0
		May,2010	10	1	0	0	0	0
		May,2010	11	0	0	0	0	0
		May,2010	12	1	0	0	0	1
		May,2010	13	1	1	0	0	0
		May,2010	14	1	0	1	0	0
		May,2010	15	1	0	0	0	0
		May,2010	16	1	0	0	0	0
		May,2010	17	1	0	0	0	0
		May,2010	18	1	0	2	0	0
		May,2010	19	1	0	0	0	ND
		May,2010	20	1	0	0	0	0
		May,2010	21	1	0	1	0	0
		May,2010	22	1	1	0	0	0
		May,2010	23	1	0	0	0	0
		May,2010	24	1	0	0	0	0
		May,2010	25	0	0	0	0	0
		May,2010	26	0	0	0	0	0
		May,2010	27	1	0	0	0	0
		May,2010	28	1	1	0	0	0
		May,2010	29	1	0	0	0	0
		May,2010	30	1	1	1	0	0
Total number of samples with GUS staining				25	7	8	0	2

CLO III-22 (pB31: pCAM 1300-35S46-GUS)	425	20-Apr	1	0	0	0	0	1
Total number of samples =36		20-Apr	2	0	0	1	1	0
		20-Apr	3	0	0	0	0	0
		20-Apr	4	DNG	DNG	DNG	DNG	DNG
		19-Jul	5	1	0	0	0	0
		19-Jul	6	1	0	0	0	0
		19-Jul	7	1	0	0	0	0
		19-Jul	8	1	0	0	0	0
		19-Jul	9	1	0	0	0	0
		19-Jul	10	1	0	0	0	0
		19-Jul	11	1	0	0	DNG	DNG
		19-Jul	12	1	0	0	0	0
		19-Jul	13	1	0	0	0	0
		19-Jul	14	0	0	0	0	0
		19-Jul	15	1	0	0	0	0
		19-Jul	16	1	0	0	0	0
		19-Jul	17	1	0	0	0	0
		06-Oct	18	2	1	1	0	0
		06-Oct	19	3	2	1	0	0
		06-Oct	20	3	1	0	0	0
		06-Oct	21	3	0	0	0	0
		06-Oct	22	3	1	0	0	0
		06-Oct	23	3	1	0	0	0
		06-Oct	24	3	1	0	0	0
		06-Oct	25	3	1	2	0	0
		06-Oct	26	2	3	0	0	1
		06-Oct	27	3	0	1	0	0
		06-Oct	28	3	1	1	0	0
		06-Oct	29	3	0	1	0	DN
		06-Oct	30	3	0	0	0	0
		06-Oct	31	0	1	0	0	0
		06-Oct	32	0	0	0	0	DN
		06-Oct	33	1	0	0	0	0
		06-Oct	34	0	1	0	0	0
		06-Oct	35	DNG	0	1	0	0
		06-Oct	36	0	0	1	0	0
Total number of samples with GUS staining				**26**	**11**	**9**	**1**	**2**
CLO III-27 (pB31: pCAM 1300-35S46-GUS)	128	Mar30.2011	1	1	0	0	0	DNG
Total number of samples =13		Mar30.2011	2	1	0	DNG	0	DNG
		Mar30.2011	3	1	1	1	1	DNG
		Mar30.2011	4	1	0	0	0	0
		Mar30.2011	5	1	0	2	0	DNG
		Mar30.2011	6	1	0	0	1	DNG
		Mar30.2011	7	2	0	3	0	0
		Mar30.2011	8	1	0	0	0	0
		Mar30.2011	9	2	0	0	0	0
		Mar30.2011	10	2	3	0	0	DNG
		Mar30.2011	11	1	0	0	0	0
		Mar30.2011	12	1	0	0	0	0
		Mar30.2011	13	2	0	0	0	0
Total number of samples with GUS staining				**13**	**2**	**3**	**2**	**0**

CLO III-43 (pB31: pCAM 1300-35S46-GUS)	151	20-Apr	1	0	0	0	0	0
Total number of samples =13		20-Apr	2	0	0	1	1	1
		20-Apr	3	0	0	0	0	0
		20-Apr	4	0	0	0	0	0
		24-Apr	5	0	0	0	1	1
		24-Apr	6	1	1	1	1	2
		24-Apr	7	0	1	0	0	0
		24-Apr	8	0	1	0	1	0
		29-Apr	9	0	0	0	0	0
		29-Apr	10	2	2	2	2	1
		02-May	11	2	2	0	0	DNG
		2011,Apri	12	0	0	0	0	1
		2011,Apri	13	0	0	0	0	0
Total number of samples with GUS staining				3	5	3	5	5
CLO III-52 (pB31: pCAM 1300-35S46-GUS)	446	02-May	1	0	0	0	0	0
Total number of samples =15		May,2011	2	2	1	1	0	0
		Jun.2011	3	0	0	0	0	0
		Jun.2011	4	0	0	0	0	0
		Jun.2011	5	0	0	0	0	0
		Jun.2011	6	0	0	0	0	0
		Jun.2011	7	0	0	0	0	0
		Jun.2011	8	0	0	0	0	0
		Jun.2011	9	1	0	0	DNG	DNG
		Jun.2011	10	0	0	0	0	DNG
		Dec,2011	11	0	0	0	0	DNG
		Dec,2011	12	0	0	0	0	DNG
		Dec,2011	13	1	0	0	0	DNG
		Dec,2011	14	0	0	0	0	DNG
		Dec,2011	15	0	0	0	DNG	DNG
Total number of samples with GUS staining				3	1	1	0	0
CLO III-53 (pB31: pCAM 1300-35S46-GUS)	152							
Total number of samples =15		2011, Apri	1	1	3	2	1	0
		2011, Apri	2	3	3	2	2	2
		2011, Apri	3	3	2	1	2	0
		2011, Apri	4	3	3	1	3	1
		2011, Apri	5	1	2	1	2	1
		2011, Apri	6	0	2	1	1	DNG
		2011, Apri	7	3	3	3	3	3
		2011, Apri	8	3	2	1	2	2
		2011, Apri	9	1	2	2	1	1
		2011, Apri	10	2	3	3	3	3
		2011, Apri	11	1	2	1	1	DNG
		2011, Apri	12	3	3	3	3	DNG
		2011, Apri	13	3	3	3	3	1
		2011, Apri	14	2	1	0	0	0
		2011, Apri	15	0	0	0	DNG	DNG
Total number of samples with GUS staining				13	14	13	13	8

CLO III-55 (pB31: pCAM 1300-35S46-GUS)	432	12.3.2010						
Total number of samples =24		26-Oct	1	1	0	0	0	0
		26-Oct	2	1	0	0	0	0
		26-Oct	3	1	0	DNG	0	0
		26-Oct	4	1	0	0	0	0
		26-Oct	5	1	0	0	DNG	0
		26-Oct	6	1	0	0	0	0
		26-Oct	7	1	1	0	0	0
		26-Oct	8	1	0	0	0	0
		Nov .25	9	1	0	0	0	0
		Nov .25	10	1	0	0	1	0
		Nov .25	11	1	0	0	0	0
		Nov .25	12	1	0	1	1	0
		Nov .25	13	1	0	0	0	0
		Nov .25	14	1	0	0	0	0
		Nov .25	15	1	0	0	0	0
		Nov .25	16	1	2	0	0	0
		Nov .25	17	1	0	0	0	0
		Nov .25	18	1	1	0	0	0
		Nov .25	19	1	1	0	0	0
		Nov .25	20	1	0	0	0	0
		Nov .25	21	1	0	0	0	0
		Nov .25	22	1	0	0	0	0
		Nov .25	23	0	0	0	0	0
			24	1	0	0	0	0
Total number of samples with GUS staining				**23**	**4**	**1**	**2**	**0**
CLO III-57 (pB31: pCAM 1300-35S46-GUS)	154	19-May	1	0	0	0	0	0
Total number of samples =16		19-May	2	0	0	0	0	0
		19-May	3	0	0	0	1	1
		19-May	4	1	1	1	2	1
		19-May	5	1	1	1	2	1
		25-May	6	0	1	0	2	2
		25-May	7	0	0	0	0	1
		26-May	8	0	0	0	2	1
		26-May	9	0	0	0	2	0
		26-May	10	0	2	2	1	1
		27-May	11	0	1	1	1	0
		27-May	12	0	1	1	1	1
		27-May	13	0	1	0	0	0
		27-May	14	0	1	0	0	0
		27-May	15	0	0	0	0	0
		27-May	16	0	0	0	0	0
Total number of samples with GUS staining				**2**	**8**	**5**	**8**	**8**
CLO III-58 (pB31: pCAM 1300-35S46-GUS)	150	Jul,2011	1	0	0	0	DNG	DNG
Total number of samples =3		Jul,2011	2	0	0	0	DNG	DNG
		Feb.2012	3	0	0	0	0	0
CLO III-63 (pB31: pCAM 1300-35S46-GUS)	154	19-May	1	0	0	0	0	0
Total number of samples =7		19-May	2	0	0	0	0	0
		19-May	3	0	0	0	3	1
		19-May	4	0	1	1	2	1
		19-May	5	0	0	0	2	1
		19-May	6	2	1	0	3	1
		19-May	7	1	1	0	2	0
Total number of samples with GUS staining				**2**	**3**	**1**	**5**	**4**

CLO III-74 (pB31: pCAM 1300-35S46-GUS)	153	Jun,2011	1	0	0	0	0	0
Total number of samples =6		Jun,2011	2	0	0	0	0	0
		Jun,2011	3	2	2	3	1	2
		Jul,2011	4	1	0	1	0	0
		Feb,2012	5	3	3	3	3	3
		Feb,2012	6	0	1	0	0	1
Total number of samples with GUS staining				3	3	3	2	3
CLO III-78 (pB31: pCAM 1300-35S46-GUS)	~450	19-May	1	0	0	0	0	0
Total number of samples =33		19-May	2	0	0	0	0	0
		19-May	3	0	0	0	0	0
		19-May	4	0	0	0	0	0
		19-Jul	5	0	0	0	0	0
		19-Jul	6	0	0	0	0	0
		19-Jul	7	0	0	0	0	0
		19-Jul	8	0	0	0	0	0
		19-Jul	9	0	0	0	0	0
		19-Jul	10	0	0	0	0	0
		19-Jul	11	0	0	0	0	0
		19-Jul	12	0	0	0	0	0
		19-Jul	13	0	0	0	0	0
		19-Jul	14	0	0	0	0	0
		19-Jul	15	0	0	0	0	0
		19-Jul	16	0	0	0	0	0
		19-Jul	17	0	0	0	0	0
		14-Oct	18	0	0	2	0	0
		14-Oct	19	1	0	1	0	1
		14-Oct	20	2	0	1	0	0
		14-Oct	21	2	0	3	0	1
		14-Oct	22	1	0	1	0	0
		14-Oct	23	2	0	0	1	1
		14-Oct	24	1	0	0	0	0
		14-Oct	25	3	0	3	0	0
		14-Oct	26	2	0	0	0	0
		14-Oct	27	1	0	0	0	0
		14-Oct	28	1	0	0	0	0
		14-Oct	29	2	0	0	0	0
		14-Oct	30	1	0	0	0	0
		14-Oct	31	1	0	0	0	0
		14-Oct	32	1	0	0	0	0
		14-Oct	33	1	0	0	0	0
Total number of samples with GUS staining				15	0	6	1	3
CLO III-80 (pB31: pCAM 1300-35S46-GUS)	451	May,2011	1	2	0	0	0	0
Total number of samples =9		May,2011	2	0	0	0	0	0
		May,2011	3	0	0	DNG	DNG	DNG
		May,2011	4	0	0	DNG	DNG	DNG
		Jun.2011	5	0	0	0	0	0
		Oct,2011	6	0	0	0	0	0
		Dec,2011	7	0	0	0	DNG	DNG
		Dec,2011	8	0	0	0	0	0
		Dec,2011	9	0	0	0	DNG	DNG
Total number of samples with GUS staining				1	0	0	0	0

Table A1: Results of GUS staining of oligonucleotide sequences cloned into pB31 plasmid vectors.

* ND= Not done because the sample died before doing it. DNG= Did Not Grow

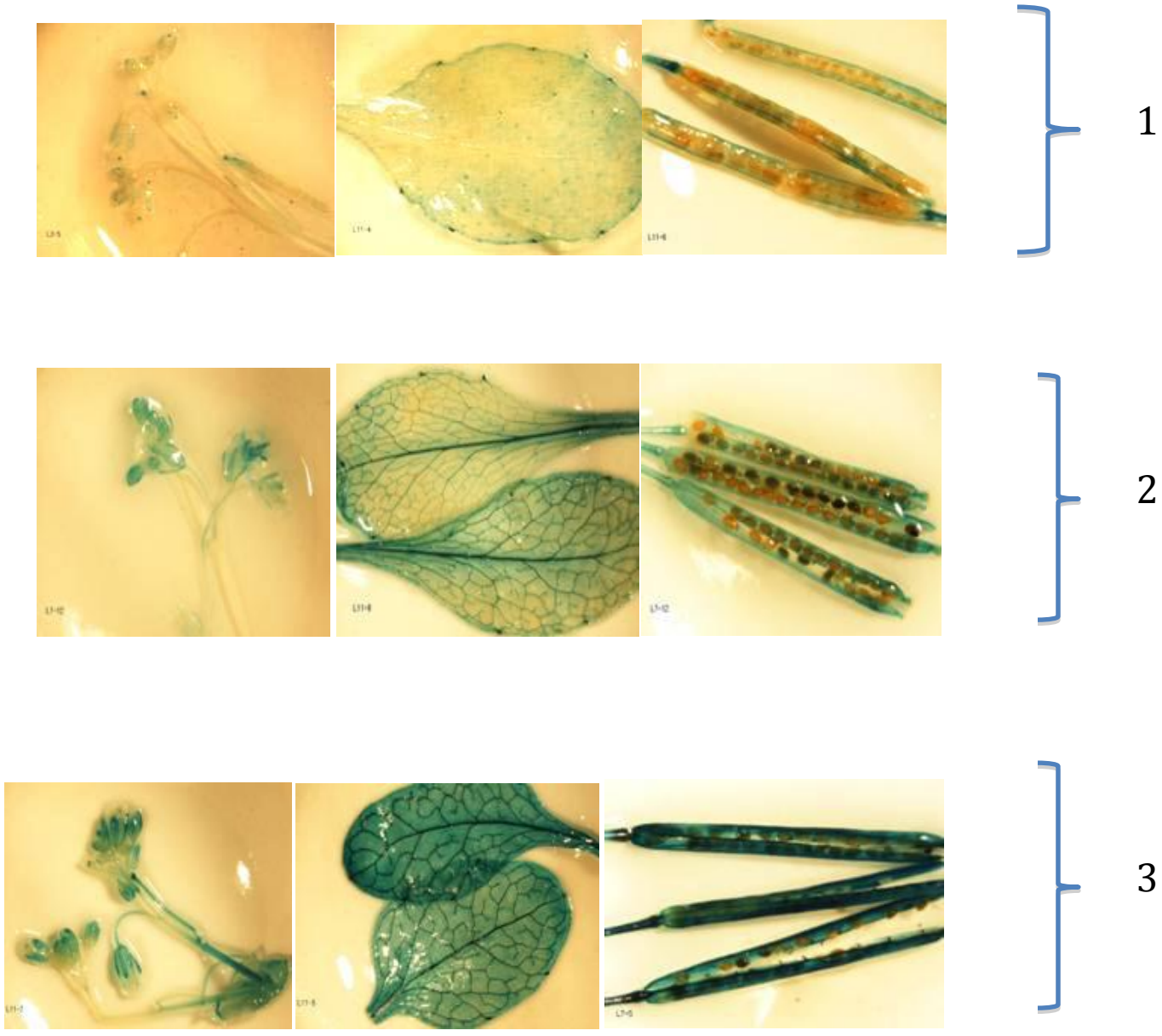


Figure A1

Scale used to assess the intensity of GUS staining. #1 means that the intensity of the blue dye in transgenic plant tissues is weak; #2 means that the strength of the blue dye in transgenic tissues is medium; #3 means transgenic tissues are strongly stained.