Direct Selection of Functional Insulators in Arabidopsis thaliana

Anh Tran
University of Ottawa Faculty of Science
Loreta Gudynaite-Savitch
University of Ottawa Faculty of Science
Lara Rasooli
University of Ottawa Faculty of Science
Douglas Allan Johnson (Johnson@Uottawa.ca)
University of Ottawa Faculty of Science https://orcid.org/0000-0003-2526-9906

Methodology

Keywords:

Posted Date: February 21st, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1336342/v1

Abstract

Background: Insulators are DNA sequences found in all eucaryotes. By forming DNA-protein complexes *in vivo* they regulate enhancer-promoter interactions and define heterochromatin-euchromatin boundaries. Studies to date have focused mainly on well-studied elements from *Saccharomyces cerevisiae*, *Homo sapiens* and *Drosophila melanogaster*. More recently there has been increased interest in insulators in plants such *Petunia hybrida* and *Arabidopsis thaliana,* in part due to the increased production of transgenic plants and there use in agriculture. The production of transgenic plants requires the correct tissue-specific and developmentally regulated gene expression of the transgene; however, their construction identified several problems such as nonspecific enhancer–promoter interactions and transgene silencing or "position effects". Insulators have the potential to mitigate these unwanted effects. Through the study of insulator mutations, we have identified several non-plant insulators that function in the model plant *A. thaliana.* For the potential of insulators to be maximized, more, different insulators need to be identified. In this paper we describe a novel protocol to isolate *de novo* insulators, enhancer-blocking DNAs that function as regulatory elements in *A. thaliana.*

Results: The selection of insulators depends upon three vectors we have constructed. pC1 is a negative selection vector. The presence of an insulator blocks the expression of the gene cytosine deaminase (*codA*) than depends upon the CaMV35S enhancer. Cytosine deaminase converts the non-toxic, cytosine analogue 5-Fluorocytosine (5-FC) into a highly toxic metabolite 5-Fluorouracil (5-FU). Plants that survive selection are potential insulators. These sequences can to cloned into pB31 that is similar to pC1 except GUS expression is monitored to confirm lack of expression in most tissues. Finally cloning into pL1 is used to verify GUS expression from the tissue-specific expression from the napin, seed-coat specific promoter is unaffected by the CaMV35S enhancer. As a proof of concept, we have recovered five, relatively-short functional insulators sequences from a library of random DNA sequences generated *in vitro*.

Conclusions: The protocol we describe can be easily scaled-up to generate more insulators. Furthermore, the source of the DNA could be any species, and modifications specific for a defined species in other genera could lead to the identification of a wide range of species-specific insulators. The elimination of unwanted enhancer-promoter interactions will impact the design of transformation systems and the recovery of transgenic plants. Now you can create your own!

Introduction

Insulators are regulatory sequences that in eucaryotes block unwanted enhancer-promoter interactions and define heterochromatin-euchromatin boundaries in a wide variety of model systems. Their function is mediated by DNA-protein interaction. For example *Drosophila melanogaster* gypsy insulator (Parkhurst et al. 1988) with the consensus sequence YRYTGCATAYYY binds the protein Suppressor of Hairy-wing Su(Hw) (Dorsett 1990; Spana and Corces 1990). Studies in the model plant *Arabidopsis thaliana* have identified DNA sequences with insulator activity based upon assays that detect enhancer-blocking activity or barrier-element activity. The first category includes, a "gypsy-like sequence" from *A. thaliana* (Singer and Cox 2013) but mainly consists of elements from other species e.g., UASrpg (Upstream Activation Site for Ribosomal Protein Genes), BEAD1c (Blocking Element Alpha/Delta-1 of the human T-cell receptor α/ δ locus), gypsy and

TBS/MAR (Transformation Booster Sequence/Matrix Attachment Region) isolated from *Ashbya gossypii*, *Homo sapiens*, *Drosophila melanogaster* gypsy retrotransposon and *Petunia hybrida* respectively (Gudynaite-Savitch et al. 2009; Tran and Johnson, 2020; Hily et al. 2009; Singer et al. 2011; Singer and Cox 2013). Tran and Johnson, 2020, went further and defined insulators within the cloned sequences by mutational analyses. The assumption that all insulators selected as enhancer-blocking elements can also function as a barrier-element (or vice-versa) is unproven. The *Petunia hybrida* TBS/MAR remains the only experimentally verified plant insulator with both barrier-element activity (Pérez-González and Caro, 2019) and enhancer-blocking activity (Singer et al., 2011). The HS4 insulator at the 5' end of the chicken β-globin locus containing a CTCF-binding site also has both activities that are separable by subcloning (Pikaart et al., 1998; Recillas-Targa et al. 2002). Based upon evolutionary arguments, the mechanism(s) underlying these two insulator activities are expected to be conserved. The expression of non-plant insulators in plants described in the literature supports this assumption.

The use of insulators flanking transgenes is a promising strategy in plant biotechnology to retaining correct regulation of the introduced gene of interest. Generating transgenic plants has identified several problems including nonspecific enhancer–promoter interactions between the promoter used to express the transgene and enhancers included with the promoter of the selective gene that interfere with the desired, correct tissue-specific, developmentally regulated gene expression. Transgene silencing or “position effects” arising from interactions between the chromosomal environment at the site of insertion that may be due to enhancers in the flanking region or result from heterochromatinization. To mitigate these problems different strategies have been investigated (Gudynaite-Savitch et al. 2009) including replacement of the CaMV 35S promoter/enhancer driving expression of the selective gene with the strong, constitutive promoter, tCUP, from tobacco (Malik et al., 2002). Misexpression can also be mitigated by adding a large segment of DNA in order to increase the distance between two promoters (Jagannath et al., 2001; Gudynaite-Savitch et al. 2009). A similar strategy, the addition of DNA segments to flank the gene(s) of interest, can be used to minimize positions effects. Both approaches would increase the size of the transgene and may decrease the recovery of transformants (Park et al., 2000). Inclusion of the *P. hybrida* TBS/MAR insulator in transgenes as flanking sequences increased expression and in one case reduced “variability between lines and between individuals” (Pérez-González and Caro, 2019). with both barrier-element activity (Pérez-González and Caro, 2019).

A wider range of insulators is needed whose functions and interactions with elements that regulate plant gene expression are well understood within the plant. Employing homologous (plant with plant) rather than heterologous (plant with non-plant) systems may also be advantageous. Based upon studies of multiple insulators in *D. melanogaster* and *in planta* experiments described previously, we hypothesize that plants may contain several different sequences with insulator function. Singer et al. (2010) unexpectedly discovered that a 1-kb bacteriophage lambda sequence may insulator function, indicating that non-plant DNA sequences can also fortuitously possess insulator function in *A. thaliana*. These observations led us to search for novel DNA sequences that can act as functional insulators in plants.

In this paper we present the results from a pilot project with the objective to isolate many different, strong insulators that function with two different promoters. To select for insulator sequences *de novo*, a three-step method was developed. Candidate insulator sequences were selected by their ability to block expression of

the selective gene c*od*A that depends upon an enhancer-promoter interaction (negative selection). Potential insulator sequences were tested in two additional vectors each with a different pattern of GUS expression from a different promoter. Our goal was to isolate many different, strong insulators that function with different promoters. These insulators may have broad applicability in the construction of transgenic plants.

Materials And Methods

Generation of a random oligonucleotide library

A random oligonucleotide library was generated in the selection vector pC1 (Fig. 1A). A collection of 124 nucleotide (nt) random sequences (5'-AGTGGATCCGAGACAAGC-(N124)-CCTCCTCCTGAATTCTGC-3' flanked by 18 nt primer binding sequences containing *Bam*H1 or *EcoR*I sites was synthesized at McMaster University, Hamilton, Canada (https://healthsci.mcmaster.ca/mobix). This collection was converted to dsDNA and amplified by PCR using the primer pair CLO-Libr-Fbpx and CLO-Libr-Resx (Table S1), digested with *Bam*HI plus *Eco*RI and ligated into pC1 which had also been digested with *Bam*HI plus *Eco*RI. Following transformation into electrocompetent *E. coli* DH10 cells (Invitrogen), plasmid DNA was purified in bulk and transformed into electrocompetent *Agrobacterium tumefaciens* GV3101 (Wang et al. 2006). Transformants were selected on agar media containing 50µg/mL kanamycin and 20µg/mL rifampicin. Details of the methodology have been described in Gandorah, 2012.

Molecular techniques

PCR reactions were performed as follows: (1) 95°C denaturation for 5 minutes, (2) 94°C for 30 seconds, (3) 55°C annealing for 30 seconds, (4) 72°C extension for 30 seconds. Steps 2–4 were repeated 31 times and a final extension step at 72°C for 5 minutes was performed to ensure proper extension of DNA. PCR products were then run on 1.25% agarose gel at 90 Volts for approximately 45 minutes to identify and/or characterize the PCR products. All PCR primers are summarized in Table S1.

Basic molecular techniques for restriction and ligation followed the manufacturers instructions. Plasmid preparation for cloning and sequencing used the GeneElute Plasmid Mini-Prep Kit (Sigma-Aldrich) or EZ-10 Plasmid DNA Minipreps Kit (Bio Basic Inc.). Cloning of PCR fragments into pGEM®-T Easy (Promega Corporation). Potential insulator sequences were amplified with the correct primer pairs to add restriction sites, digested with the correct pair of restrictions enzymes and ligated into the target vector digested with the same pair of enzymes. For cloning of sequences from pC1 into pB31 primer pair pCLO-Libr-Fbpx and CLO-Libr-Resx was used with the restriction enzymes *Bam*HI plus *Eco*RI. For cloning sequences from pB31 into pL1 the same primer pair was used with the restriction enzymes *Hin*dIII plus *Pst*I. Latterly the Takara Bio USA, Inc In-fusion® HD cloning technique was used following manufacturer's instructions and online primer design protocols
(http://www.clontech.com/US/Products/Cloning_and_Competent_Cells/Cloning_Resources/Online_In-Fusion_Tools).

Transgenic *A. thaliana* genomic DNA for PCR and/cloning was extracted using a small-scale procedure (McKinney et al., 1995) with minor modifications.

Plant materials and DNA analyses

Mature *A. thaliana* wild type plants underwent *Agrobacterium*-mediated transformation using the floral dip method (Clough & Bent, 1998). Seeds were then collected and grown for selection on Murashige and Skoog medium (Murashige and Skoog, 1962) containing 30µg/ml hygromycin-B for selection of plants carrying the T-DNA insert and 250µg/ml timentin to inhibit bacterial growth, augmented with 2,000 µg /ml 5-Fluorocytosine to inhibit the growth of plants lacking a functional insulator (Supplementary Information, "Screening for 5-Fluorocytosine Resistance"). Seeds were grown on selective media until shoots and roots were observed, transferred to a second selective plate for up to 3 weeks and finally transferred to soil. After 3–4 weeks samples of flowers, leaves and siliques were collected into 1.5ml microtubes with a second sample of siliques collected one week later for samples cloned into pL1 (pCAMBIA 1391 napin-GUS). A sample would contain 3–4 organs from the sample transgenic plant. GUS histochemical staining was carried out as described by (Malik et al. 2002). Scoring was done independently by two persons and was stringent - if any part of any replicate within a tube stained blue, the sample was scored as positive for staining (Tran and Johnson 2020).

To verify the transgenic DNA sequences, PCR products were generated from each transgenic was extracted using a small-scale procedure (McKinney et al., 1995) with minor modifications. DNAs were sequenced at Plant Biotechnology Institute, Saskatoon, Saskatchewan) or McGill University and Genome Quebec Innovation Centre or the Ottawa Hospital Research Institute. Sequences for InI-3 (OM273459), InII-10 (OM273460), InII-12 (OM273461), InIII-50 (OM273462), InIII-78Δ5'(OM273463) have been deposited in GenBank™.

Bioinformatics

DNAMAN (http://www.lynnon.com/) and Sequencher DNA Sequence Analysis (https://www.genecodes.com/) were employed for the analysis of DNA sequencing results and chromatographs. For searches and comparisons, programmes at European Bioinformatics Institute (https://www.ebi.ac.uk/) were used.

Results

Strategy for Insulator Selection

Insulator isolation was initially based on a 3-step process in which vectors were designed for the selection of candidate insulator (In) sequences, the testing of re-isolated sequences in a second vector to confirm insulator activity and final cloning into a third vector to determine that the sequence was not a strong repressor and that expression from a tissue-specific seed coat promoter was observed.

pC1 (pCAM1300-35S46-*codA*) is a negative selection vector (Fig. 1A) containing cytosine deaminase gene (*codA*) (Perera et al., 1993) expressed from the core CaMV35S promoter (35S46) (Covey et al., 1981). Expression is boosted by CaMV35S enhancer ("35S double enhancer") responsible for driving expression of hygromycin phosphotransferase II (hptII) used for plant selection. Cytosine deaminase converts the non-toxic, cytosine analogue 5-Fluorocytosine (5-FC) into a highly toxic metabolite 5-Fluorouracil (5-FU) (1989; Perera et al., 1993; Stougaard, 1993) that *in planta* irreversibly inhibits thymidylate synthase thus depriving cells of

dTTP and leading to cell death (Mullen et al., 1992) The pathway is illustrated in Fig, S1. The presence of a DNA fragment with a functional insulator between the CaMV35S enhancer and 35S46 will block enhancer influence on *codA* and plants will survive. Loss of *codA* expression for one of many possible reasons such as cloning or an insulator or a strong repressor or mutation within *codA* that could allow plant survival. Hence the need to further testing.

pB31 (pCAMBIA1300-35S46-GUS) is a screening vector with in which *codA* from pC1 is replaced by the β-glucuronidase (GUS) reporter gene (Fig. 1B). As in pC1, the presence of an insulator between the CaMV35S enhancer and the 35S46 core promoter will block the expression of the transgene resulting in no blue stain in any of the tissues tested -flowers, leaves, siliques- while the absence of a functional insulator will result in constitutive GUS throughout the plant. This does not distinguish between an insulator and a strong repressor.

The final vector, pL1 (pCAMBIA 1391 napin-GUS) uses a similar strategy to pB31 except that GUS is now expressed from the napin seed-specific promoter (Ericson et al., 1991) in lieu of the 35S46 core promoter (Fig. 1C). In pL1 the napin promoter under the influence of the CaMV35S enhancer drives non-specific GUS expression, showing blue staining in all tissues tested-flowers, leaves, siliques. Cloning of a strong repressor would block all gene expression, resulting in no staining in transgenic tissues (Gudynaite-Savitch et al. 2007). An insulator cloned between the two elements blocks non-specific expression, as a result GUS expression is found only in seeds and at the same time verifies that the insulator functions with two different promoters.

Key: hptII: hygromycin phosphotransferase; codA: cytosine deaminase; GUS: marker gene β-glucuronidase; 35S: CaMV 35S promoter S46: CaMV 35S core promoter Napin: *Brassica napus* seed-specific promoter

Selection of insulator sequences

124 bp random DNA oligonucleotide library plus flanking sequences was cloned into pC1 for transformation into *A. thaliana*. Plantlets containing potential insulator sequences were selected on agar plants containing 30µg/ml hygromycin-B and 250µg/ml timentin supplemented with 2000 µg /ml 5-Fluorocytosine. Initial screening identified 100 plantlets that survived selection on agar plates; subsequent transfer to soil resulted in 60 candidate plants with putative insulators. Seed was collected from these plants for further propagation and DNA was isolated from leaf tissue for PCR and sequence analyses. They were classified into four groups (I, II, III, and IV) based upon the selection trial.

Twenty-one were chosen based on plant growth and health on selective media and in soil for subsequent cloning of the insert (potential insulator) into the vector pB31. Upon further analysis, nine had the phenotype expected for a fragment without insulator activity (Table 1A), eight had the phenotype expected for insulator activity (Table 1B) and four were in the category “too few transgenic plants to infer the category”.

These results suggest that the 5-FC can be used to select insulator sequences from a library of random sequences. However, a strong repressor activity could be expected to mimic an insulator in this assay. To differentiate between these two possibilities inserts from clones in Table 1B were cloned into the vector pL1 in which an insulator positioned between the 35S double-enhancers and the napin promoter will block the action of the enhancers but not the tissue-specific napin promoter expressing GUS (Gudynaite-Savitch et al. 2009). Clone InIII-80 may also be a good candidate for an insulator but at this stage too few transgenic plants were

analysed (Table 1B). As indicated in Table 1C, 5 candidates, InI-3, InII-10, InII-12, InIII-50, and InIII-78Δ5' (a spontaneous deletion of InIII-78) demonstrate the expected insulator GUS phenotype. Based upon this small survey, we conclude that insulator activities can be isolated *de novo* from a random library. Actually, the source of the DNA should be immaterial.

Defining Insulator Sequences

The sizes of the inserts presented in Table 1, and in the other clones not yet completely analysed, are not just the expected 154 bp for the insert plus primer binding sites but range from ~ 150 bp up to 446 bp. Sequencing revealed that the ~ 450 bp inserts were ligation products of three different ~ 154 bp sequences to give a trimeric insert (Gandorah, 2012). The minor differences between them could result from mutations/deletions generated during PCR and prior to ligation or by deletions due to recombination during plant transformation.

For further bioinformatics approaches the longer sequences were subdivided into 3 units (A,B,C) *in silico* based upon the positions of the internal *BamH*I and *EcoR*I sites. Further editing removed the primer binding sites used for PCR and the terminal restriction sites used for cloning from each unit. Treated in this way resulted in a total of 11 sequences were recovered for the insulators InI-3, InII-10, InII-12, InIII-50, and InIII-78Δ5'. None were identical, suggesting we had identified at least 5 unique insulator sequences.

Several insulators from plant or other sources have been defined (Pérez-González and Caro 2019; She et al. 2010; Tran and Johnson 2020). Our initially focus was on the group UASrpg, BEAD1c, gypsy to determine whether our experimentally defined insulators contained sequences that were similar to them. Both orientations were investigated. Finally, a set of insulators with known sequences (scs, CoreScs, GAGA, MCP, HS4, HS185) was tested. No 100% matches were observed but near identical matches were identified (Fig. S3); InI-3 with, MCPRC (7/9) and CoreScsRC (8/10); InII-10 with UASrpg1RC (8/10); InII-12 with UASrpg1 (7/10) and UASrpg1RC (7/10); InIII-50 with scsRC (7/10). No matches were observed for InIII-78_5'.

1A. GUS staining results for candidate sequences cloned into pB31 with no apparent insulator activity. The expected phenotype of insulator activity is no GUS expression in all tissues.

			GUS positive staining				GUS expression
		No. plants tested					
Construct Name	Size* (bp)		Flowers	Leaves	Siliques	Stained	Not stained
pB31		62	11	39	38	74.2%	25.8%
InII-3	126	17	9	6	9	76.5%	23.5%
InII-7	128	30	11	14	8	63.3%	36.7%
InIII-4	414	26	16	4	3	65.4%	34.6%
InIII-17	417	30	25	15	2	86.7%	13.3%
InIII-22	399	36	26	11	3	80.6%	19.4%
InIII-27	102	13	13	4	2	100.0%	0.0%
InIII-53	126	15	13	14	14	93.3%	6.6%
InIII-55	406	23	22	5	2	95.7%	4.3%
InIII-57	128	16	2	8	9	68.8%	31.2%

3B. GUS staining results for candidate sequences cloned into pB31 with apparent insulator activity. The phenotype of insulator activity is no GUS expression in all tissues.

			GUS positive staining			GUS expression	
		No. plants tested					
Construct Name	Size* (bp)		Flowers	Leaves	Siliques	Stained	Not stained
InI-3	412	51	0	0	0	0.0%	100.0%
InI-6	436	24	0	0	0	0.0%	100.0%
InII-10	128	40	0	0	0	0.0%	100.0%
InII-12	401	37	0	0	3	8.1%	91.9%
InIII-50	417	11	0	0	0	0.0%	100.0%
InIII-52	420	15	3	1	0	20.0%	80.0%
InIII-78	407	20	0	0	0	0.0%	100.0%
InIII-80	425	9	1	0	0	11.1%	88.9%

1C. GUS staining results for candidate sequences re-cloned into pL1 with apparent insulator activity. The phenotype of insulator activity is GUS expression only in seeds.

Table 1

The β-glucoronidase (GUS) staining results for potential *Arabidopsis thaliana* insulator sequences. The columns indicate the sequence number, construct name, size, number of total transformations, the number of individual transformants tested for flower, leaf and silique staining, and a calculation of the number of stained and unstained samples. The data sets for plant staining and potential insulator sequences can be found in Supplementary Information Table S2.1 (Staining of pB31 transgenic plants), Table S3.2 (Staining of pL1 transgenic plants) and Fig. S2 (Verification of insulator sequences in transgenic pL1)

			GUS positive staining				**Seed-specific GUS expression**		
Construct Name	Size* (bp)	No. plants tested	Flowers	Leaves	Siliques	Seeds	Specific	Nonspecific	None
pL1		53	41	49	42	41	3.9%	96.1%	2
InI-3	412	20	0	0	2	16	89.8%	11.1%	2
InII-10	128	27	0	1	0	22	95.7%	4.4%	4
InII-12	401	14	0	0	0	10	100.0%	0.0%	4
InIII-50	417	13	0	0	0	12	100.0%	0.0%	0
InIII-78Δ5’	116	16	0	0	1	14	93.3%	6.7%	1

* The size of the insert cloned into either pB31 or pL1 after removal of the primer sequences.

Discussion

We have successfully recovered several novel insulator sequences from a random sequence library that are functional in *A. thaliana.* They range in size from 116 bp to 417 bp. Each has insulator activity when test in two vectors, inserted into pB31 between the CaMV35S enhancer driving expression of hptII and the CaMV35S core promoter (35S46) driving expression of GUS in pB31 or between the CaMV35S enhancer and the napin promoter in pL1. And in each case non-specific enhancer-promoter interaction is almost completely blocked.

Simple analyses of their sequences (Fig. S3) suggests that some contain sites similar to insulators in other species. These comparisons were performed between sequences from different species and some differences may reflect evolutionary divergence. When no comparisons were observed, such sequences may represent newly identified insulators. Further genetic analysis is needed to identify the precise sequences required for insulator activity (Tran and Johnson, 2020).

This was a pilot study in *A. thaliana* where several candidates were identified by screening a small number of plants from selective plates containing 5FU. The selective system is scalable, thus there is scope for the

isolation of even more candidates using the same wide-spread technologies. The essential regulatory elements derived from CaMV function in many plant species. Napin, originally from *Brassica napus*, has a more limited range but it should be straightforward to substitute another tissue-specific promoter that functions in the species of interest. Modifications of the original strategy should also improve insulator recovery. For example, given the high percentage of insulators recovered, it should be possible to omit cloning into pB31 and using fusion technology to clone directly into pL1 and using restriction enzymes.

Declarations

Ethics approval and consent to participate: None required.

Consent for publication: None Required.

Availability of data and materials: Sequences have been submitted to GenBank and accession numbers are included in manuscript. The sequences are included in Supplementary Information.

Competing interests: None

Funding: This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada (RGPIN/ 002869-2011) and the University of Ottawa (555256) to DAJ.

Authors' contributions: A.T. L.G.S and D.A.J. conceived and designed the research. All authors conducted experiments, analyzed data, and wrote the manuscript. All authors read and approved the manuscript.

Acknowledgements

Several undergraduate and honours students helped with collection of data including Shukria Ahmadi, Hassan Badreddine, Onkar Bhanushaki, Linda Dam, Krishna Gelda Ka Mien Lam, Liam McCarthy, Adina Popescu and Tatjana Semiz. We thank Dr. Ethan R. Signer, MIT, for sending us the *cod*A gene.

References

1. Bi X, Broach JR. UAS(rpg) can function as a heterochromatin boundary element in yeast. Genes Dev. 1999;13:1089–101. https://doi.org/10.1101/gad.13.9.1089.
2. Dorsett D. Potentiation of a polyadenylylation site by a downstream protein-DNA interaction. Proc Natl Acad Sci U S A. 1990;87:4373–7. https://doi.org/10.1073/pnas.87.11.4373.
3. Ericson ML, Muren E, Gustavsson HO, et al. Analysis of the promoter region of napin genes from brassica-napus demonstrates binding of nuclear-protein invitro to a conserved sequence motif. Eur J Biochem. 1991;197:741–6. https://doi.org/10.1111/j.1432-1033.1991.tb15966.x.
4. Gandorah B. (2012) Identifying Insulators in Arabidopsis thaliana Master of Science thesis https://ruor.uottawa.ca/bitstream/10393/23226/1/Gandorah_Batool_2012_thesis.pdf.
5. Gudynaite-Savitch L, Johnson DA, Miki BLA. Strategies to mitigate transgene-promoter interactions. Plant Biotechnol J. 2009;7:472–85. https://doi.org/10.1111/j.1467-7652.2009.00416.x.

6. Hark AT, Schoenherr CJ, Katz DJ, et al. CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. Nature. 2000;405:486–9.
7. Higo K, Ugawa Y, Iwamoto M, Korenaga T. Plant cis-acting regulatory DNA elements (PLACE) database: 1999. Nucleic Acids Res. 1999;27:297–300. https://doi.org/10.1093/nar/27.1.297.
8. Hily J-M, Singer SD, Yang Y, Liu Z. A transformation booster sequence (TBS) from Petunia hybrida functions as an enhancer-blocking insulator in Arabidopsis thaliana. Plant Cell Rep. 2009;28:1095–104. https://doi.org/10.1007/s00299-009-0700-8.
9. Jagannath A, Bandyopadhyay P, Arumugam N, Gupta V, Burma PK, Pental D. The use of a spacer DNA fragment insulates the tissue-specific expression of a cytotoxic gene (barnase) and allows high-frequency generation of transgenic male sterile lines in Brassica juncea L. Mol Breed. 2001;8:11–23.
10. Malik K, Wu K, Li X-Q, Martin-Heller T, Hu M, Foster E, Tian L, Wang C, Ward K, Jordan M, Brown D, Gleddie S, Simmonds D, Zheng S, Simmonds J, Miki B. A constitutive gene expression system derived from the tCUP cryptic promoter elements. Theor Appl Genet. 2002;105:505–14.
11. McKinney EC, Ali N, Traut A, et al (1995) Sequence-based identification of T-DNA insertion mutations in Arabidopsis-actin mutants Act2-1 and Act4-1. Plant J 8:613–22. https://doi.org/10.1046/j.1365-313X.1995.80406 13.x.
12. Murashige T, Skoog F. (1962) A revised medium for rapid growth and bio assays with Tobacco tissue cultures. Physiol Plant 15:473–97. https://doi.org/10.1111/j.1399-3054.1962.tb080 52.x.
13. Park SH, Lee B-M, Salas MG, Srivatanakul M, Smith RH. Shorter T-DNA or additional virulence genes improve Agrobacterium-mediated transformation. Theor Appl Genet. 2000;101:1015–20.
14. Parkhurst SM, Harrison DA, Remington MP, et al. The Drosophila Su(Hw) gene, which controls the phenotypic effect of the gypsy transposable element, encodes a putative DNA-binding protein. Genes Dev. 1988;2:1205–15. https://doi.org/10.1101/gad.2.10.1205.
15. Pérez-González A, Caro E. Benefits of using genomic insulators flanking transgenes to increase expression and avoid positional effects. Sci Rep. 2019;9:8474. https://doi.org/10.1038/s41598-019-44836-6.
16. Perera RJ, Linard CG, Signer ER. Cytosine deaminase as a negative selective marker for Arabidopsis. Plant Mol Biol. 1993;23:793–9.
17. Pikaart MJ, Recillas-Targa F, Felsenfeld G. Loss of transcriptional activity of a transgene is accompanied by DNA methylation and histone deacetylation and is prevented by insulators. Genes Dev. 1998;12:2852–62. doi:10.1101/gad.12.18.2852.
18. Rasooli L. (2014) Analysis of Potential DNA Insulators in Arabidopsis thaliana. Master of Science thesis https://ruor.uottawa.ca/bitstream/10393/31329/1/Rasooli_Lara_2014_thesis.pdf.
19. Recillas-Targa F, Pikaart MF, Burgess-Beusse B, Bel AC, Litt MD, West AG, Gaszner M, Felsenfeld G. Position-effect protection and enhancer blocking by the chicken β-globin insulator are separable activities. Proc Natl Acad Sci. 2002;99(10):6883–8. .
20. She W, Lin W, Zhu Y, et al. The gypsy Insulator of Drosophila melanogaster, Together With Its Binding Protein Suppressor of Hairy-Wing, Facilitate High and Precise Expression of Transgenes in Arabidopsis thaliana. Genetics. 2010;185:1141-U27. https://doi.org/10.1534/genetics.110.117960.

21. Singer SD, Cox KD. A gypsy-like sequence from Arabidopsis thaliana exhibits enhancer-blocking activity in transgenic plants. J Plant Biochem Biotechnol. 2013;22:35–42. https://doi.org/10.1007/s13562-012-0108-3.
22. Singer SD, Hily J-M, Cox KD. Analysis of the enhancer-blocking function of the TBS element from Petunia hybrida in transgenic *Arabidopsis thaliana* and *Nicotiana tabacum*. Plant Cell Rep. 2011;30:2013–25. https://doi.org/10.1007/s00299-011-1109-8.
23. Singer SD, Hily J-M, Liu Z. A 1-kb Bacteriophage Lambda Fragment Functions as an Insulator to Effectively Block Enhancer-Promoter Interactions in Arabidopsis thaliana. Plant Mol Biology Report. 2010;28:69–76. https://doi.org/10.1007/s11105-009-0122-3.
24. Spana C, Corces VG. DNA bending is a determinant of Binding-specificity for a Drosophila zinc finger protein. Genes Dev. 1990;4:1505–15. https://doi.org/10.1101/gad.4.9.1505.
25. Stougaard J. Substrate-dependent negative selection in plants using a bacterial cytosine deaminase gene. Plant J. 1993;3(5):755–61. doi:10.1111/j.1365-313x.1993.00755.x.
26. Tran A. (2018) Investigation of Enhancer-Blocking DNA Insulators in Arabidopsis thaliana https://ruor.uottawa.ca/handle/10393/37829.
27. Tran A, Johnson DA. (2020) Mutational analysis identifies functional Rap1, Su(Hw), and CTCF insulator sites in Arabidopsis thaliana. Plant Cell Reports PCRE-D-20-00325R1. https://rdcu.be/b7wMR.
28. Wang WY, Esch JJ, Shiu SH, et al. Identification of important regions for ethylene binding and signaling in the transmembrane domain of the ETR1 ethylene receptor of Arabidopsis. Plant Cell. 2006;18:3429–42. https://doi.org/10.1105/tpc.106.04453.
29. Yu Q, Qiu RX, Foland TB, et al. Rap1p and other transcriptional regulators can function in defining distinct domains of gene expression. Nucleic Acids Res. 2003;31:1224–33. https://doi.org/10.1093/nar/gkg200.

Figures

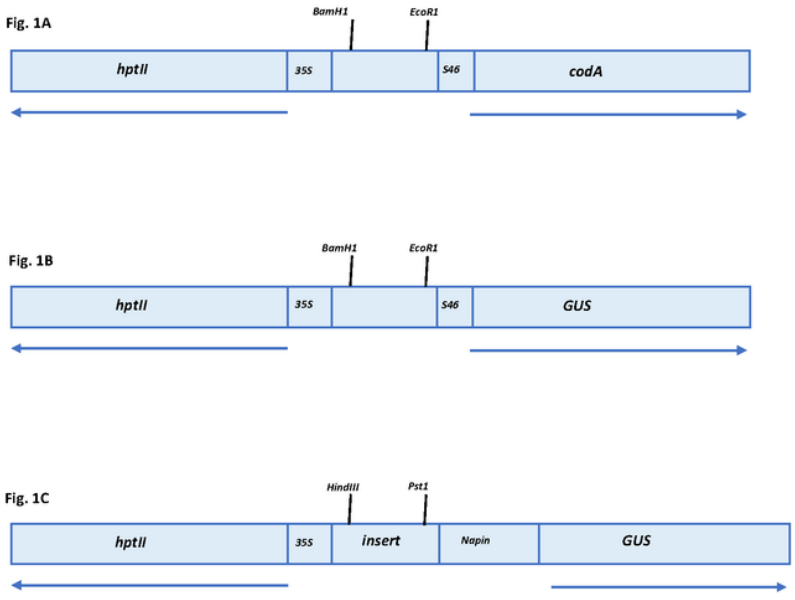

Figure 1

Schematic representation of T-DNA inserts used for selection and screening of potential insulator sequences. **(Fig. 1A)** Vector pC1(pCAM1300-35S46-*CodA*); used for the isolation of potential insulator sequences by negative selection using the *CodA* gene (blue box). **(Fig. 1B)** Vector pB31 (pCAM1300-35S46-GUS); used to confirm the presence of a potential insulator sequence using the GUS gene for screening. **(Fig. 1 C)** Vector pL1 used to confirm the presence of a potential insulator sequence using the GUS gene for screening and to test that insulator function is not promoter-dependent. Adapted from Tran (2018). Not drawn to scale.

Key: hptII: hygromycin phosphotransferase; codA: cytosine deaminase; GUS: marker gene β-glucuronidase; 35S: CaMV 35S promoter S46: CaMV 35S core promoter Napin: *Brassica napus* seed-specific promoter

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- InsulatorDiscoverySupplementalFInal.docx