Improved T-DNA vector for tagging plant promoters via high-throughput luciferase screening

Serge Remy, Els Thiry, Bert Coemans, Saskia Windelinckx, Rony Swennen, and László Sági

BioTechniques 38:763-770 (May 2005)

Transferred DNA (T-DNA) tagging is a powerful tool for tagging and in planta characterization of plant genes on a genome-wide scale. An improved promoter tagging vector is described here, which contains the codon-optimized luciferase (\textit{luc}+) reporter gene 31 bp from the right border of the T-DNA. Compared to the wild-type luciferase gene, this construct provides significantly increased reporter gene expression and a 40 times higher tagging frequency. The utility of the construct is demonstrated in banana, a tropical monocot species, by screening embryogenic cell colonies and regenerated plants with an ultrasensitive charged-coupled device (CCD) camera. The improved vector resulted in a luciferase activation frequency of 2.5% in 19,000 cell colonies screened. Detailed molecular analysis of flanking DNA sequences in a tagged line revealed insertion of the luciferase tag in a novel gene with near-constitutive expression.

INTRODUCTION

Insertional mutagenesis for promoter tagging has become a powerful tool for functional genomics, because it allows the genome-wide identification and isolation of functionally redundant genes that would not be possible by conventional mutagenesis approaches. The characteristics of the tagged genes can then be studied in their native context throughout development and in various tissues. The characterized promoters could also be utilized in transgenic product development, when specifically regulated expression of the genes of interest is required.

Promoter tagging in plants is traditionally based on using a promoterless selectable or screenable (reporter) gene, which is linked to a transferred DNA (T-DNA) border, and after integration is activated by flanking promoters (1). Besides genome size, the efficiency of plant transformation and the sensitivity of the reporter gene used are the major factors that determine the efficiency of promoter tagging. At present, the most sensitive reporter gene system is based on the firefly luciferase (\textit{luc}) gene, which also allows monitoring temporal changes in gene expression. Initially, this system has therefore been used to study circadian regulation of plant gene transcription (2), but its application to promoter and gene tagging in higher plants has also been reported recently (3–5).

Here we describe an improved promoter tagging vector based on a codon-optimized \textit{luc}+ gene and demonstrate its use and increased tagging frequency in banana (\textit{Musa} spp.), a tropical monocot plant.

MATERIALS AND METHODS

Plant Material

Embryogenic cell suspensions of the banana cultivar Three Hand Planty (accession no. of International Transit Centre ITC.0185) were maintained and subcultured in ZZ medium (6), which is half-strength MS medium (7) supplemented with 5 \(\mu\)M 2,4-D and 1 \(\mu\)M zeatin.

Tagging Vectors

For construction of pETKUL2, a promoterless codon-optimized \textit{luc}+ gene (http://www.promega.com/pnotes/49/2788c/2788C_core.pdf; GenBank® accession no. U47122) was inserted close to the right T-DNA border of the binary vector pCAMBIA2300. A 2-kb \textit{luc}+\textit{-nos} fragment was amplified from plasmid VpFAJ3204 (received from M. De Bolle, Centre for Microbial and Plant Genetics, Katholieke Universiteit Leuven) using the primers Pluc+L and PtnosR (all primer sequences are given in Table 1). Following digestion with \textit{PmeI}, the PCR product was ligated into the \textit{PmeI} linearized pCAMBIA2300 vector (GenBank accession no. AF234315, position 8622). In the resulting pETKUL2 plasmid (10.7 kb) the start codon of the \textit{luc}+ gene is located 31 bp from the right T-DNA border. To construct the control vector pETKUL3, primers B35SL and Blucl+R.
containing BstBI sites were used in PCR with VpFA3204 as template. The resulting 1-kb PCR product containing the enhanced cauliflower mosaic virus (CaMV) 35S RNA promoter with tobacco mosaic virus leader sequence and the first 172 bp of the luc+ gene was digested with BstBI and ligated into the BstBI-digested pETKUL2 plasmid. The latter digestion removed the first 172 bp of the luc+ gene. As a result, in plasmid pETKUL3 the luc+ gene is controlled by the enhanced CaMV 35S promoter (Figure 1A). Both constructs also contain the neo gene fused to the promoter and terminator regions of the CaMV 35S RNA (Figure 1A). The vectors and corresponding sequence information are available on request.

Plasmids were transferred to One Shot® TOP10 chemically competent Escherichia coli (Invitrogen, Merelbeke, Belgium) by heat shock and to Agrobacterium EHA105 (8) by electroporation. The integrity of the vectors was verified by custom DNA sequencing (Flanders Institute for Biotechnology, Genetic Research Facility, Antwerp, Belgium).

**Agrobacterium-Mediated Transformation**

Agrobacterium transformation of embryogenic suspension cultures was done as described previously (9) with some modifications. Briefly, agrobacteria were grown in YEP medium (10 g/L Bacto yeast extract, 10 g/L Bacto peptone, 5 g/L NaCl, pH 7.5) supplemented with the appropriate antibiotics for 20–24 h. Bacterial cultures were diluted to an A600 of approximately 0.4 units and then transferred to ZZ medium containing 200 µM acetosyringone (AS) for induction. Each sample of embryogenic suspension cells containing 200 µL at 33% settled cell volume (±50 mg fresh weight of cells) was mixed with 1 mL induced agrobacteria in a well of a 24-well plate (Greiner Bio-One, Wemmel, Belgium) and incubated in the dark at 25°C on a rotary shaker at 25 rpm for 6 h.

Following 1 week of cocultivation on solid ZZ medium supplemented with 200 µM AS, cells were transferred to Petri plates with solid selective ZZ medium containing 50 mg/L genetin and 200 µg/L timentin and subcultured every 2 weeks for 2 to 3 months. Transgenic colonies were picked up individually and transferred for regeneration to 24-well plates containing solid proliferation and regeneration medium as described (10).

**Screening for Luciferase Activation**

Screening for luciferase (LUC) activation was regularly performed during in vitro culture and regeneration. First, 2 to 3 months after transformation, 40 µL luciferin solution (0.1 mM in half-strength MS medium) per sample were applied twice to embryogenic cell cultures on Petri plates at an intermittent period of 24 h before picking positive colonies. Second, transgenic colonies after selection were rescreened 1 month later in 24-well plates containing solid ZZ medium by saturation with the luciferin solution mentioned above. Third, differentiating shoots, in vitro plants, and their bisected meristems, as well as leaf discs of 5 × 5 cm and root pieces excised from greenhouse-grown plants were placed horizontally on water-saturated filter paper in a Petri dish and sprayed twice (24 h in between) with 200 µL luciferin solution using a type 250-2 airbrush (Badger Airbrush Company, Franklin Park, IL, USA). Plantlets were wounded by several scalpel incisions in the pseudostem and the leaves to improve penetration of the substrate solution.

Light emission was captured in a light-tight box (Progress Control, Waalwijk, The Netherlands) with a liquid nitrogen-cooled Versarray™ 512 B LN charge-coupled device (CCD) camera (Roper Scientific B.V., Vianen, The Netherlands) attached to a light-sensitive camera lens (Nikkor F 50 mm f/1.2; Nikon, Tokyo, Japan). A live reference image was first taken, and LUC images were then recorded for 20 min after an incubation of a minimum of 3 h in the dark and processed using the image analysis software MetaMorph 5.0r3 (Universal Imaging, Downingtown, PA, USA). LUC images for control and pETKUL3 cell colonies were processed with a high-scale setting of 600 (i.e., the number of different grey levels), whereas those for pETKUL2 to a high-scale setting of 2000. All plant LUC images were scaled to a high setting of 5000.

**Southern Blot Hybridization**

Total DNA isolation and digestion, gel electrophoresis, blotting, and hybridization were performed as previously described (10). Labeled luc- and cloned sequence-specific probes were produced by standard PCR using digoxigenin-11-dUTP. Immunochromiluminescent detection of bound probe using the CSPD® substrate (Roche, Vilvoorde, Belgium) was
done according to the instructions of the manufacturer. Chemiluminescent light emission was captured with the liquid nitrogen-cooled CCD camera as described above, using an integration time of up to 20 min.

**Thermal Asymmetric Interlaced PCR**

Genomic sequences flanking either the right or left T-DNA border in tagged cultures, and plants were amplified by the standard PCR cycling program (11) using the 128-fold degenerate primer AD2 (5′-NGTCGASWGANAWGAA-3′) (11) and T-DNA border-specific nested primers. For the right border, these primers were TAILRBLUC1, TAILRBLUC2, and LUC-R3 in the primary, secondary, and tertiary reaction, respectively; for the left border, the corresponding primers were TAILLbpET2n1, TAILLbpET2n4, and TAILLbpET2n3.

**PCR and Reverse Transcription**

**PCR Analysis**

For PCR analysis, the 5′-tagged sequence-specific primers ET2-344-S1 and ET2-344-S2 were used in combination with ET2-344-R2 (these primers are depicted—see Figure 3A), which is specific for the tagged downstream coding region. The calculated product sizes were 629 and 246 bp, respectively.

Total RNA was isolated by the Aurum Total RNA Mini kit (Bio-Rad, Nazareth Eke, Belgium) followed by an extra DNase treatment in Mn buffer (12). cDNA was synthetized by the Omniscript® RT kit (Qiagen, Venlo, The Netherlands). The 5′-tagged sequence-specific primer ET2-344-S2 or ET2-344-S3 was used in combination with LUC-R3 (see Figure 3A) to yield an expected product of 150 or 126 bp, respectively, as deduced from the sequenced thermal asymmetric interlaced PCR (TAIL-PCR) product. A banana actin gene served as positive control and to confirm the absence of genomic DNA (12).

**Sequencing and Sequence Analysis**

TAIL-PCR and reverse transcription PCR (RT-PCR) products were cloned in pCR4-TOPO® vector (Invitrogen) and commercially sequenced. Analysis of the tagged downstream sequence was performed using Basic Local Alignment Search Tool (BLAST, blastn, and blastx; www.ncbi.nlm.nih.gov/BLAST) and FASTA (www.ebi.ac.uk/fasta33) programs in the GenBank, The Institute for Genomic Research (TIGR; www.tigr.org/tdb/e2k1/osa1), and the Rice Genome Research Program (RGP; riceblast.dna.affrc.go.jp) databases, and by searching the Interpro database (Reference 13; www.ebi.ac.uk/interpro). The sequences were matched with banana expressed sequence tags (ESTs) in a database donated by Syngenta to the Global Musa Genomics Consortium. The 5′-tagged sequence was analyzed by searching the PLACE database (Reference 14; www.dna.affrc.go.jp/PLACE) and the PlantCARE database (Reference 15; intra.psb.ugent.be:8080/PlantCARE).

**RESULTS**

**Large-Scale Luciferase Screening Using the Improved Tagging Vector**

The tagging vector pluc19 has been previously described as a means for tagging plant promoter sequences. This construct contains a promoterless wild-type luciferase (luc) gene 662 bp from the left T-DNA border and the neomycin phosphotransferase (neo) selectable marker gene fused to the nopaline synthase (nos) promoter and followed by the nos 3′ poly(A) region (Figure 1A) (16). To assess its frequency of tagging promoter sequences in banana, the pluc19 construct was first introduced with Agrobacterium into embryogenic suspension cells. Since plant regeneration takes more than 6 months, screening for LUC activation was initiated 2 to 3 months after transformation in cell colonies without induction treatment. Among the several
hundreds of independent transformed cell colonies per plate, usually one cell colony (Figure 1B) or none displayed baseline luciferase activation (BLA). Furthermore, detected luminescence was generally weak (Figure 1B), which required careful identification and isolation of the tagged cell colonies. These initial findings were confirmed on a large scale in two independent experiments (Table 2). By screening more than 26,000 independent transformed cell colonies, a BLA frequency of 0.06%–0.07% was obtained. Since an ultrasensitive luminescent detection system was used (17), we assumed that most putative tagged colonies with BLA were captured and, consequently, that the observed tagging frequencies well represented the actual ones. Using the same construct, a 50- to 100-fold higher BLA frequency of 3.3% was obtained in tobacco shoots (3). As the banana genome is significantly smaller than that of tobacco (approximately $0.6 \times 10^9$ bp versus $1.5 \times 10^9$ bp), we concluded that promoter tagging in banana could be hampered by low expression levels of the wild-type luc gene and/or a low frequency of intact insertions.

Therefore, the following three modifications were made to construct the tagging vector pETKUL2 (Figure 1A). First, the codon-optimized $\text{luc}^+$ gene was used, since previous transient and stable LUC expression assays had demonstrated about 20 times higher activation frequency in bombarded banana cells than the wild-type luc gene when both genes were driven by the maize ubiquitin promoter (17) and light intensity was significantly higher in transgenic luc$^+$ plants both in vitro and in the greenhouse (data not shown). Second, as T-DNA strand excision is initiated at the right T-DNA border and integration is likely to result in less rearrangement at this border (18,19), the promoterless luc$^+$ gene was placed near the right T-DNA border. Third, the distance between the border sequence and the luc$^+$ start codon was reduced from 662 bp in pluc19 to 31 bp in pETKUL2 in order to enhance the likelihood of LUC activation.

An increased tagging frequency was achieved with pETKUL2 even by simple visual inspection, and on average, five or more cell colonies per sample showed BLA (Figure 1B). As could be anticipated from the above, LUC expression was also stronger in cell colonies transformed with pETKUL2 than with pluc19 (Figure 1B), which facilitates the isolation of promoter-tagged cell colonies and demonstrates the importance of engineering the reporter gene.

Table 2. Baseline Luciferase Activation in Transformed Banana Cell Colonies

<table>
<thead>
<tr>
<th></th>
<th>Experiment 1</th>
<th>Experiment 2</th>
<th>Experiment 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total No. Cell Colonies Screened</td>
<td>24,540</td>
<td>1550</td>
<td>19,000</td>
</tr>
<tr>
<td>BLA Frequency (%)</td>
<td>0.07</td>
<td>0.06</td>
<td>2.50</td>
</tr>
</tbody>
</table>

Frequencies were determined 2 to 3 months after Agrobacterium-mediated transformation.

*aBaseline luciferase activation (BLA) frequency—the number of cell colonies showing luciferase activity without induction expressed in percentage of the total number of independent transformed cell colonies screened.
addition, many tagged colonies showed an activity comparable to or higher than those containing the enhanced CaMV 35S RNA promoter (Figure 1B, pETKUL3).

Figure 2 represents a range of luc+ expression levels and patterns in a transgenic banana population tagged with pETKUL2. Screening at the colony level has so far resulted in 51 colonies with, on average, 170 times higher expression levels (range: 10–400 times) compared to that obtained in 24 colonies transformed with the control vector pETKUL3 (Figure 2A). Twenty-nine of these 51 tagged colonies (56%) were regenerated, of which 25 lines (86%) showed high levels of activated expression in whole in vitro plants, well above the background and the level that was obtained with the 35S promoter control in only 3 out of 17 plants (Figure 2B). This near-constitutive expression pattern has been characterized and is shown as an example in two tagged lines (Figure 2C, ET2-171 and ET2-344). Both lines possess a significantly higher expression than the CaMV 35S promoter-driven control (pETKUL3) in cell cultures and in vitro plants. A comparison of the two lines reveals higher expression level in ET2-171 during the undifferentiated cell culture period, whereas ET2-344 exerts higher activity in several tissues of regenerated plants (Figure 2C).

As can be seen in Table 2, the BLA frequency in cell colonies reached up to 2.5% with pETKUL2, a 40-fold increase compared to the 0.06% obtained with pluc19. Similarly high activation frequencies were recently obtained with pKC1, a vector that carries an intron-containing luc+ gene in the same vector backbone (data not shown).

Tagging Near-Constitutive Promoters

Tagged line ET2-344 showed high LUC activity in all tissues of in vitro plants, including the meristematic region (Figure 2C). Moreover, a strong LUC activity was detectable in mature leaves and roots of a greenhouse-grown plant from this line (data not shown), indicating that a near-constitutive promoter sequence has been tagged.

The flanking genomic DNA in both directions was isolated by TAIL-PCR using primer sets specific either to the right border (Figure 3B, lanes 1 and 2) or the left border (Figure 3B, lanes 3 and 4) region of the inserted T-DNA. Based on sequencing the 546-bp TAIL-PCR product downstream of the left border in ET2-344, the T-DNA was found integrated 60 bp upstream of the start codon of a banana gene. Over the length of the 486-bp coding sequence isolated so far, this gene shows high homology to several rice cDNAs (GenBank accession nos. AK067663 and AK069166; both at E < 6 × 10^{-77} as well as to the first 135 amino acids of an unknown protein (NP_922270; 80% identity and 89% positives at E < 5 × 10^{-61}) encoded by the first exon of a rice gene (NM_197288) on chromosome 10. The additional high homology to a 384-bp banana EST (600096752T1; 90% identity at
E < $1 \times 10^{-75}$) in the second half of this sequence (between 298 and 539 bp) indicates that it is fully transcribed. However, no conspicuous functions or protein domains could be identified for any of these sequences.

Analysis of the 739-bp sequence in the 5′-tagged sequence revealed the presence of two potential TATA box regions within a region 200–350 bp upstream of the start codon.

Genomic Southern hybridizations demonstrated two luc+ inserts in line ET2-344 (Figure 3C, lane 2) as already indicated by TAIL-PCR analysis (Figure 3B). A comparison on the same blot between the hybridization patterns obtained with the luc probe and the 5′-tagged sequence-specific probe (Figure 3C, lanes 4 and 5) allowed the identification of a common fragment. This confirms the physical linkage of the cloned 5′ region with one of the luc+ inserts.

To find out whether it is the cloned 5′ region that activated the luc+ reporter gene, RT-PCR was performed with either of two primers binding within a distance of 25 bp in this region in combination with a luc+ specific primer (Figure 3, A and D). The presence and the correct sequence of the expected products with the diagnostic 25-bp length shift in line ET2-344 only (Figure 3D, lanes 1 and 2) confirms that the tagged 5′ sequence can indeed be associated with the transcription of the activated luc+ gene.

Finally, PCR analysis with primers specific to the 5′-tagged sequence in combination with a downstream sequence-specific primer (Figure 3, A and E) resulted in specific products with the calculated length both in an untransformed control plant and in line ET2-344 (Figure 3E, lanes 1 and 4, and lanes 2 and 5). This result indicates that the cloned 5′ region upstream of the right T-DNA border and the 3′ sequence downstream of the left border form a continuous sequence.

Line ET2-344 thus represents a tagged event, in which the luc+ reporter gene has marked a new and highly active gene in banana. The presented data also demonstrate the utility of the new pETKUL2 vector for tagging and the functional characterization of (novel) plant genes.

**DISCUSSION**

We have constructed a binary vector, pETKUL2, based on a codon-optimized luciferase (luc+) reporter gene, for tagging genes and their promoters in plants via Agrobacterium-mediated transformation. In this vector, the promoterless luc+ gene was placed only 31 bp from the right T-DNA border in order to increase the frequency of tagging. Indeed, activation frequency

![Figure 3. Molecular characterization of tagged banana line ET2-344.](image)
as measured by luciferase luminescence in transformed cell colonies was increased 40-fold, from 0.06% with pluc19 (a vector containing a wild-type luc gene 662 bp from the left T-DNA border) (16) to 2.5% with pETKUL2.

This activation frequency is comparable to that reported with the promoterless wild-type luc gene in tobacco (3.3%) (3) and in Arabidopsis (3.7%) (5) and is also similar to the frequency observed with a promoterless uidA reporter gene in Arabidopsis (5%) (20) and tobacco (6%) (21). Using a T-DNA tagging construct carrying the uidA gene near the right T-DNA border, a β-glucuronidase (GUS) expression frequency of 0.9% and 1.1% was achieved in rice for leaves and roots, respectively (22). It should be noted that screening for reporter gene activation starts at an earlier stage in the banana system than in the model crops where it is done at the regenerated plant level, making direct comparisons in tagging frequency difficult. Taking into consideration that one person is able to screen full time about 50,000 independent transformed cell colonies per week, the early screening method presented here can be a useful alternative to in planta screening for certain applications. In this case, one has to take into consideration that after transformation, in vitro selection of banana cell colonies requires 2 months, whereas plant regeneration and line production takes another 4 months. Importantly, the new tagging construct can also be used in a wide range of monocots and especially cereals, because the luc+ gene has been efficiently expressed by particle bombardment in maize and wheat cells (23).

Using the pETKUL2 vector, we have been able to select transgenic lines (such as ET2-171; Figure 2C) with highly activated expression preferentially during the undifferentiated growth stage. Possibly, some of the corresponding promoters can be used to direct selectable marker gene expression in cell colonies, which might be more desirable from a public perception viewpoint than using for instance a plant virus promoter for the same purpose.

An important point for the evaluation of a gene tagging system is the question of whether the observed expression pattern and activity of the reporter gene corresponds to that of the tagged native gene. A correlation between the reporter gene activity and the transcription of the endogenous (wild-type) gene largely depends on the exact position and integration context of the tag. Therefore, in many tagged lines it will be difficult to find a precise correlation. We will study this question by comparing expression levels of selected tagged and native genes via real-time PCR. A further confirmation of the promoter characteristics requires an additional back-transformation of the cloned sequences, which is in progress for a number of tagged genes.

Another factor that may interfere with the interpretation of the results is the potential survival of agrobacteria in selected cell colonies. Background expression from these contaminating bacteria could lead to the overestimation of activation frequencies and to picking false positive colonies. The survival of agrobacteria is not limited to in vitro tissues and plants but also reported in greenhouse plants (24) and thus may, in principle, present a problem even for the in planta transformation of Arabidopsis. In our case, previous control PCR experiments with primers specific to the picA gene (25), which resides on the bacterial chromosome (26), were never positive for transgenic colonies or plants (data not shown), indicating no survival of agrobacteria in the tagged lines. This conclusion is further supported by another line of evidence coming from tagging experiments with the pKC1 vector (not shown), which contains the same T-DNA as pETKUL2 but carries an intron in the luc+ gene. Use of the pKC1 vector resulted in similar activation frequencies (between 1%–3%) as obtained with pETKUL2. Since, as can be expected, the same intron in the wild-type luc gene was not functional in Agrobacterium (27), the similar range of activation frequency in the case of the pKC1 and pETKUL2 vectors can be attributed to the integrated luc+ gene rather than to contaminating bacteria.

In conclusion, we have described an improved tagging vector for high-throughput luciferase-based promoter and gene tagging in banana. This tagging vector could be used in other monocots, for which an efficient trans-
formation protocol exists, and possibly in a broader range of organisms, which can be transformed by Agrobacterium (28–30). An interesting application of the presented screening technology could be the search for inducible genes, because homogeneous cell cultures can be easily and reliably induced, for example, by temperature or a chemical treatment.

ACKNOWLEDGMENTS

This work was supported by a grant of the Belgian Administration for Development Cooperation (BADC-GOS). The pluc19 vector was provided by Robert Birch (University of Queensland, Brisbane, Australia). The authors thank three anonymous reviewers for useful comments.

COMPETING INTERESTS STATEMENT

The authors declare no competing interests.

REFERENCES


Received 23 December 2003; accepted 22 December 2004.

Address correspondence to:
Serge Remy
Laboratory of Tropical Crop Improvement
Katholieke Universiteit Leuven
Kasteelpark Arenberg 13
3001 Leuven, Belgium
E-mail: serge.remy@biw.kuleuven.be