

Somnath Bhattacharyya · Sitakanta Pattanaik  
Indu B. Maiti

## Intron-mediated enhancement of gene expression in transgenic plants using chimeric constructs composed of the *Peanut chlorotic streak virus* (PCISV) promoter–leader and the antisense orientation of PCISV ORF VII (p7R)

Received: 5 November 2002 / Accepted: 14 June 2003 / Published online: 22 July 2003  
© Springer-Verlag 2003

**Abstract** The antisense orientation of the *Peanut chlorotic streak virus* (PCISV) open reading frame (ORF) - VII (denoted as p7R), in conjunction with the sense orientation of the PCISV leader sequence, acts as an intron and enhances the expression of a reporter gene, analyzed in protoplasts and transgenic plants of tobacco (*Nicotiana tabacum* L.). Correct 5' and 3' splicing sites were determined for intron removal from the chimeric constructs using either  $\beta$ -glucuronidase (GUS) or chloramphenicol acetyltransferase (CAT) as a reporter gene. In this splicing process, the active consensus 5' splicing donor site (AG/GTATA) is located at position +283 to +289 from the transcription start site (TSS) of the PCISV full-length transcript (FLt). The 3' splice site (TAG/GATT) is located on the p7R sequence at position +785 to +791 from the TSS. The combination of PCISV FLt leader and p7R enhanced the expression of reporter genes (CAT and GUS) by as much as 2-fold compared to the strong constitutive PCISV FLt promoter without an interfering leader sequence and about 30- to 800-fold compared to constructs containing the sense orientation of PCISV ORF VII (p7) in both protoplast transient-expression experiments and stably transformed transgenic plants. An increased level of mature transcripts accompanied this. This suggests that this combination of elements can mediate the intron-mediated enhancement (IME) phenomenon. We also demonstrated comparative IME with other heterologous promoters from caulimoviruses.

**Keywords** Enhanced expression · Intron · *Nicotiana* · Peanut chlorotic streak virus · Splicing · Transgenic tobacco

**Abbreviations** CaMV: cauliflower mosaic virus · CAT: chloramphenicol acetyltransferase · FLt: full-length transcript · GUS:  $\beta$ -glucuronidase · IME: intron-mediated enhancement · nt: nucleotide(s) · ORF: open reading frame · PCISV: peanut chlorotic streak virus · Sgt: subgenomic transcript · TSS: translation start site

### Introduction

The ability of natural introns to enhance gene expression has been well documented in various organisms, including mammals (Buchman and Berg 1988; Chung and Perry 1989), insects (Meredith and Storti 1993), nematodes (Okkema et al. 1993), and plants (Callis et al. 1987; Luehrsen and Walbot 1991; Rose and Last 1997). The mechanism of intron-mediated enhancement (IME), however, is not well understood. Evidence has accumulated to suggest that after transcription, introns increase the level of mRNA by enhancing the maturation and stability of nascent transcripts (Callis et al. 1987; Mascarenhas et al. 1990; Clancy et al. 1994; Donath et al. 1995; Dean et al. 1989; Rose and Last 1997). Unlike transcriptional enhancers, the enhancing effect of introns is position and orientation dependent (Callis et al. 1987; Mascarenhas et al. 1990; Snowden et al. 1996). Deletion mapping of maize introns *Adh1* and *Sh1* shows that most intron sequences are dispensable (Clancy et al. 1994; Luehrsen and Walbot 1994a) for IME; however, a large internal deletion of the first intron of the maize *Adh* gene reduces splicing and consequently IME (Luehrsen and Walbot 1994b). A point mutation at the 5'- or 3'- splice site of the maize Hsp82 intron impaired splicing, suggesting that IME requires intron splicing (Sinibaldi and Mettler 1992). In contrast,

S. Bhattacharyya · S. Pattanaik · I. B. Maiti (✉)  
Molecular Plant Virology and Plant Genetic Engineering  
Laboratory, Kentucky Tobacco Research and  
Development Center, University of Kentucky, Lexington,  
KY 40546-0236, USA  
E-mail: imaiti@uky.edu  
Fax: +1-859-3231077

*Present address:* S. Bhattacharyya  
Department of Genetics,  
Bidhan Chandra Krishi Viswavidyalaya,  
B/8-48 Kalyani, Nadia, West Bengal, India

in dicots the *PAT1* intron has been shown to increase mRNA accumulation independent of a unique intron sequence or splicing (Rose and Beliakoff 2000). The functional mechanism for the recognition and efficiency of IME differs in monocot and dicot species (Koziel et al. 1996). A variety of intron-containing fragments fused with a heterologous promoter demonstrate IME (Vain et al. 1996; Dugdale et al. 2001). However, to retain strong promoter activity, monocot-derived promoters like the maize *adh1* promoter (Dennis et al. 1984; Callis et al. 1987), the maize *ubi1* promoter (Christensen et al. 1992; Christensen and Quail 1996), the maize *shrunken-1* promoter (Maas et al. 1991) and the rice actin promoter (McElroy et al. 1990) require their native first intron to be present in the 5' untranslated region (UTR).

Peanut chlorotic streak virus (PCISV) is a species of the genus *Caulimovirus*, a subgroup of the Caulimoviridae family (review: Shepherd 1989; Rothnie et al. 1994). Cauliflower mosaic virus (CaMV) is the type species of the genus *Caulimovirus*. Like other caulimoviruses, the PCISV genome of 8,174 bp (GenBank accession no. U13988) produces two primary transcripts: one is a full-length transcript (FLt) equivalent to CaMV 35S RNA and the other is a subgenomic transcript (Sgt) equivalent to CaMV 19S RNA (Richins 1993). The PCISV FLt RNA is composed of a 346-nucleotide-long 5'-untranslated leader (5'-UTR) sequence with four small open reading frames (ORFs) followed by eight tightly arranged longer ORFs (VII-I-A-B-C-IV-V-VI). It is a complex, highly structured polycistronic RNA that encodes all the viral proteins (Richins 1993). The exact mechanism involved in translation of genes on the full-length transcript of caulimoviruses is not clearly established yet but several studies suggest that the FLt of caulimoviruses functions as a polycistronic messenger RNA.

In this report we have demonstrated that the antisense orientation of the PCISV ORF VII (denoted as p7R) in conjunction with the PCISV leader acts as an intron when included into the transcription unit of chimeric gene constructs, and enhances expression of the reporter gene, as shown both in protoplast transient-expression experiments and in stably transformed transgenic plants. The enhanced expression is due to higher accumulation of mature mRNA. In addition, we have shown that the p7R IME of expression is not gene dependent; however, this enhancement depends in part on the type of promoter sequence used. Although the physiological significance of this IME is not clear and IME is rather unlikely to participate in the virus' life cycle, the usefulness of this viral sequence, p7R, as a tool for IME is apparent.

## Materials and methods

### Protoplast isolation, electroporation and plant transformation

Isolation of protoplasts from cell suspension cultures of tobacco (*Nicotiana tabacum* L. cv. Xanthi 'Brad') and electroporation of tobacco protoplasts with supercoiled DNA were done essentially as

previously described (Maiti et al. 1998; Dey and Maiti 1999). All constructs were tested in at least four independent experiments. Constructs developed for expression in transgenic plants were introduced into *Agrobacterium tumefaciens* strain C58C1: pGV3850 by triparental mating. Tobacco plants (*N. tabacum* cv. Samsun NN) were transformed with the engineered *Agrobacterium* as described earlier (Maiti et al. 1993). Twelve independent plant lines were generated for each construct. Regenerated kanamycin-resistant plants were grown under greenhouse conditions. R1 lines with  $\text{Kan}^R:\text{Kan}^S = 3:1$  segregation were selected for further analysis.

### Chloramphenicol acetyltransferase (CAT) assay, $\beta$ -glucuronidase (GUS) assay and luciferase assay

CAT activity was determined according to Gorman et al. (1982). Plant tissue extracts containing 5  $\mu\text{g}$  of soluble protein were used for each CAT assay. The reaction was carried out at 37 °C for 30 min. The rates of reaction were linear over the period of incubation. Fluorometric assays to measure GUS activity in plant tissue or protoplast extracts were performed according to Jefferson et al. (1987) as described earlier (Maiti et al. 1997). Protein in plant extracts was determined according to the method of Bradford (1976) using BSA as a standard. Luciferase activity was measured in transfected protoplasts using a luciferase assay system (Promega, Cat. No. E1500) with a luminometer (Model No. TD2020; Turner Designs, Sunnyvale, CA, USA).

### Analysis of transgenic plants: isolation of RNA, northern blot analysis

Total cellular RNA from electroporated protoplasts or 4-week-old seedlings was isolated using the RNeasy plant mini kit (Qiagen, Chatsworth, CA, USA) and the RNA blots were done as described earlier (Bhattacharyya et al. 2002). DNA probes were generated using a random labeling kit (Prime-It; Stratagene, La Jolla, CA, USA) with [ $^{32}\text{P}$ ]dCTP. Radio-labeled probes were purified by using QIA quick nucleotide removal kit (Qiagen).

The RNA blots were probed with  $^{32}\text{P}$ -labeled DNA probes of GUS, CAT, PCISV leader, 3'rbcsE9 terminator sequence or gene VII sequence to detect the different spliced and full-length transcripts.

### Reverse transcription–polymerase chain reaction (RT–PCR) analysis

DNase-treated RNA (2  $\mu\text{g}$ ) was used in a one-tube DNase/RT reaction (20  $\mu\text{l}$ ) following the procedure described by Huang et al. (1996) with slight modifications using an RT–PCR kit (Life Technologies, Rockville, MD, USA). For the no-reverse-transcriptase control, an individual reaction was performed in parallel without addition of reverse transcriptase. A 1/20 volume (1  $\mu\text{l}$ ) of RT reaction was used in the subsequent PCR reaction with appropriately designed forward and reverse primers to detect the spliced and full-length transcripts. The PCR reaction was performed in a total volume of 25  $\mu\text{l}$  for 30 cycles (92 °C for 30 s, 55 °C for 30 s, 68 °C for 2 min). As a negative control, each primer pair was tested against DNase-treated RNA to confirm cDNA dependence of amplification. PCR products were displayed on an ethidium bromide-stained agarose gel.

### Real-time quantitative RT–PCR (qRT–PCR)

The expression level of GUS mRNA in transgenic plants developed for plasmids pKPG, pKPLG, pKPL7RG and pKPL7G was evaluated by real-time qRT–PCR as described earlier (Maiti et al. 2003). Isolation of total cellular RNA from plants and synthesis of first-strand cDNA was performed as described earlier. A 90-bp fragment was amplified from the full-length GUS transcript using

specific primer pairs. A plasmid containing the full-length GUS cDNA was used as an external control. Serial dilution ( $10^3$  to  $10^8$  copies/ $\mu$ l) of the control plasmid was used to generate a standard curve. A housekeeping gene (tobacco  $\alpha$ -tubulin, accession no. AJ421413) was used as an internal control to correct any variation in samples. PCR amplification was performed in the DNA Engine Opticon2 System for continuous fluorescence detection (MJ Research, Watertown, MA, USA) in a total volume of 20  $\mu$ l containing 1  $\mu$ l of cDNA, gene-specific primers using DyNamo SYBR Green qPCR kit (MJ Research). Each PCR reaction was performed in triplicate using the following conditions: 94 °C for 30 s; 55 °C for 30 s; 72 °C for 30 s; 78 °C for 1 s; plate read, 35 cycles followed by 5 min extension at 72 °C. Copy number of the target samples was calculated using the Opticon Monitor software. Melting curve analysis (Ririe et al. 1997) was done to characterize the amplified products by slowly raising the temperature (0.2 °C/s) from 65 °C to 95 °C with fluorescence data collected at 0.2 °C intervals.

#### Construction of plasmids pUP9, pKP9, pUP9L and pKP9L

A 377-bp fragment of the FLt promoter of PCISV (−353 to +24 from the TSS; corresponding PCISV genomic coordinates 5725–6101) was PCR-amplified as an *EcoRI*–*HindIII* fragment from the genomic clone: pPCISV-K1 (Ducasse and Shepherd 1995) using appropriately designed forward and reverse primers. The *EcoRI*–*HindIII* promoter fragment (called P9) was cloned into the corresponding sites of pUCPMA (Dey and Maiti 1999) and pKYLX71 (Schardl et al. 1987) by substituting the CaMV 35S promoter. The resulting plasmids were designated as pUP9 and pKP9, respectively. The ‘K’ in a construct name indicates a plasmid capable of stable transformation.

A 698-bp fragment of the PCISV FLt promoter with its extended leader sequence (PCISV genomic coordinates 5725–6422) was amplified as an *EcoRI*–*HindIII* fragment by PCR using pPCISV-K1 as template with appropriately designed primers. The *EcoRI*–*HindIII* promoter–leader fragment was cloned into the corresponding sites of pUCPMA (Dey and Maiti 1999) and pKYLX71 (Schardl et al. 1987). The resulting plasmids were named pUP9L and pKP9L, respectively.

#### Construction of pBp7 and pBp7R

The ORF VII of PCISV (coordinates 6424–6855 in the PCISV genome) was amplified from the genomic clone pPCISV-K1 (Ducasse and Shepherd 1995) with appropriately designed forward and reverse primers to insert an *XhoI* site at the 5′-end and *SalI*–*SstI* sites at the 3′-end. The amplified fragment was cloned as *XhoI*–*SstI* into the corresponding sites of pBS(KS+) (Stratagene). The resulting plasmid was designated as pBp7. The ORF VII was PCR-amplified with appropriately designed primers to insert *SstI*–*SalI* sites at the 5′-end and an *XhoI* site at the 3′-end, and it was cloned into the *XhoI*–*SstI* sites of pBS(KS+) (Stratagene) to create the plasmid pBp7R that contains the antisense of PCISV ORF VII.

#### Construction of plasmids pBG, pPG, pPLG, pKPG, pKPLG, pBC, pPC, pPLC, pKPC, pKPLC and pBLuc, pPLuc, pPLLuc

The coding sequence of the reporter gene (GUS, CAT or luciferase) was PCR-amplified using appropriately designed primers to insert an *XhoI* site at the 5′-end and *SalI*–*SstI* sites at the 3′-end. The PCR-amplified fragment of either GUS, CAT or luciferase was cloned separately as *XhoI*–*SstI* into the corresponding sites of pBS(KS+) (Stratagene), pUP9, pUP9L, pKP9, pKP9L to generate the plasmids pBG, pPG, pPLG, pKPG, pKPLG with the GUS reporter gene; pBC, pPC, pPLC, pKPC, and pKPLC with the CAT reporter gene; and pBLuc, pPLuc, pPLLuc with the luciferase gene, respectively.

#### Construction of pBp7G and pBp7RG; pBp7C and pBp7RC

The GUS reporter gene isolated as an *XhoI*–*SstI* fragment from the plasmid pBG was inserted into the *SalI*–*SstI* sites of pBp7 and pBp7R to create the plasmids pBp7G and pBp7RG, respectively. Similarly, the CAT reporter gene was isolated from the plasmid pBC as an *XhoI*–*SstI* fragment and cloned into the *SalI*–*SstI* sites of pBp7 and pBp7R to generate the plasmids pBp7C and pBp7RC, respectively.

#### Construction of plant expression vectors containing the P9 promoter and P9 promoter–leader sequence for transient protoplast expression and expression in stably transformed transgenic plants

The *XhoI*–*SstI* fragments from pBp7G, pBp7RG, pBp7C, and pBp7RC, were cloned separately into the corresponding sites of the protoplast expression vector pUP9 (containing the P9 promoter) and pUP9L (containing the P9 promoter–leader), to create the following plasmids: pP7G, pP7RG, pP7C, and pP7RC, respectively with the P9 promoter; and plasmids pPL7G, pPL7RG, pPL7C, and pPL7RC, respectively with the P9 promoter–leader.

For stable transformation in transgenic plants, the *XhoI*–*SstI* fragments mentioned above were cloned individually into the plant expression vector pKP9L with the promoter–leader. The following plant constructs were created: plasmids pKPL7G, pKPL7RG, pKPL7C, and pKPL7RC, respectively with the P9 promoter–leader. A control plasmid vector that lacks the P9 promoter–leader sequence was developed by *EcoRI*–*HindIII* digestion of pUP9L followed by end-filling and ligation to create the plasmid pUA. All plant expression vectors generated for this project contain the 3′*rbcsE9* terminator sequence.

#### Cloning of the spliced fragment of RT–PCR products into protoplast expression vectors

The RT–PCR products from matured spliced transcripts of KPL7RG-plants and KPL7RC-plants were produced using appropriately designed forward primer corresponding to the 5′-end of the leader (+24 from TSS) with a *HindIII* site and reverse primer corresponding to the 3′-end of the GUS or CAT gene with a *SstI* site. The PCR-amplified 5′-*HindIII*–*SstI*-3′ fragment of the partial deleted leader either with a GUS or CAT gene was cloned separately into the *HindIII*/*SstI* sites of the protoplast-expression vector pUP9 to generate the plasmid pPrtpcL<sub>284</sub>G or pPrtpcL<sub>284</sub>C. As a control, the PCISV leader sequence (coordinates +25 to +284 from TSS) was amplified as a 5′-*HindIII*–*XhoI*-3′ fragment and cloned into the corresponding sites of pPG to generate the construct pPL<sub>284</sub>G.

#### DNA sequencing of RT–PCR products

For DNA sequencing, RT–PCR products of total RNA from plants developed for constructs pKPLC, pKPL7C, pKPL7RC, pKPL7RG, pKPLG and pKPL7G were used. The following primer pairs, forward primer: 5′PL6078xh#428 (5′-GCGGGCTC-GAG ACCCGATCGAGAAGACAC-3′), 18-nucleotide (18-nt) leader sequence +1 to +18 from TSS (underlined), *XhoI* site (bold); and reverse primer: 3′*rbcsE9*/16P#444 (5′-ATGCAGCTG-CAG TGTCGAAACCGATGTTAC-3′), corresponding to the 5′-terminal 18-nt sequence of the *rbcs* polyA tract (underlined), *PstI* site (bold) were used. Automated DNA sequencing was done with an Applied Biosystems ABI Prism 310 Genetic Analyzed (Perkin Elmer) using ABI Prism Dye terminator cycle sequencing ready reaction kit containing Ampli Taq DNA polymerase, as described earlier (Dey and Maiti 1999).

## Results

Enhanced expression of genes in plants using the PCISV leader and the antisense orientation of PCISV ORF VII (p7R)

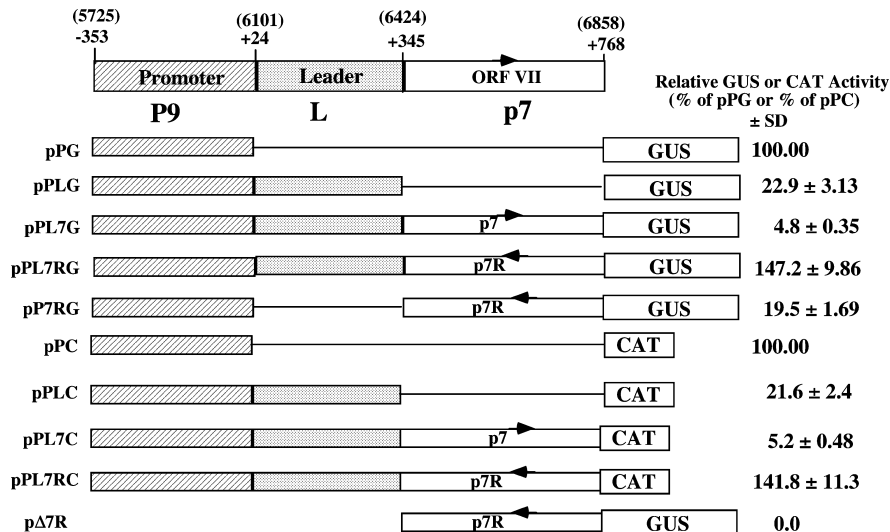
There remains some ambiguity regarding the *cis*-functional role of gene VII of caulimoviruses (Futterer and Hohn 1991; Richins 1993). To determine the effect of PCISV ORF VII (denoted as p7) on the expression of the downstream gene(s), several constructs were generated with sense and antisense orientation of p7 fused in frame with the GUS or CAT reporter gene. These constructs all possessed the P9 promoter (PCISV FLt promoter), with or without the associated leader (L). The results of protoplast transient expression experiments with these constructs are shown in Fig. 1. The expression levels seen with construct pPG or pPC, where GUS or CAT is controlled by the PCISV FLt promoter, were set as 100%. GUS activity was reduced to 22% when the leader was included (construct pPLG) and was further

reduced to 5% when the ORF VII was included along with the leader (construct pPL7G; Fig. 1) compared to construct pPG where GUS is under the strong P9 promoter. Similarly, CAT activity in construct pPLC was reduced to  $\approx 22\%$  compared to pPC (Fig. 1).

To determine the requirement of ORF VII (p7) in expression of downstream genes, p7 was replaced with the antisense orientation p7R (construct pPL7RG and pPL7RC). Surprisingly, the construct pPL7RG showed 30-fold higher GUS expression than that obtained from the construct pPL7G. The GUS expression level with pPL7RG was even 1.5-fold higher than that obtained with pPG, where GUS is under the strong constitutive PCISV FLt promoter without interfering leader sequences. Similarly, pPL7RC gave 27-fold more CAT activity than did pPL7C (Fig. 1). Our results suggest that, in this context, enhanced expression of downstream genes is not a gene-dependent phenomenon.

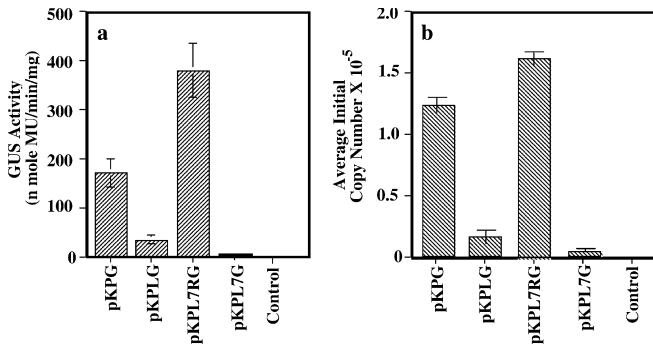
These results may reflect a promoter activity of the antisense orientation of the PCISV ORF VII coding region. However, no GUS activity was detected with the construct p $\Delta$ 7RG, where the PCISV promoter–leader region was deleted and the GUS gene is under p7R (Fig. 1). Therefore, the behavior of neither the PCISV leader nor the p7R sequence on expression of downstream genes in plants is due to promoter effects.

**Fig. 1** Effect of PCISV leader, ORF VII and antisense-ORF VII on the transient expression of downstream reporter genes (either GUS or CAT) in protoplasts of tobacco (*Nicotiana tabacum* L. cv. Xanthi). *Top* A portion of the PCISV genome (nucleotide positions 5725–6858 in parenthesis; corresponding coordinates –335 to +768 from the TSS) representing the PCISV FLt promoter (P9; –353 to +24), untranslated leader sequence (L; shaded box) and ORF VII (open box) is shown. *Below* Schematic maps of respective constructs containing the PCISV FLt promoter (P9), leader (L), sense (p7) or antisense (p7R) strand of ORF VII, and GUS (G) or CAT (C) gene are presented (boxes). Right-handed or left-handed arrowhead indicates orientation of ORF VII as p7 or p7R, respectively. Protoplast transient expression analysis was performed as described in Materials and methods. Expression of P9 luciferase was used as an internal control for transfection efficiency. The levels of GUS and CAT expression are presented as percent activity of pPG or pPC (Maiti and Shepherd 1998), and these were routinely used as positive GUS and CAT expression plasmids for these experiments. The data are the mean GUS or CAT activities ( $\pm$  SD) of five independent experiments for each construct



### Transgenic plant expression analysis

Ten to twelve independent transgenic lines were developed for each of the following GUS and CAT constructs: pKPG, pKPLG, pKL7G, pKPL7RG, pKPC, pKPLC, pKL7C, and pKPL7RC as described in Materials and methods. Integrity of the construct in transgenic plants (Ro lines) was checked by PCR and RT-PCR analysis followed by DNA sequencing (data not presented). The expression of the GUS reporter genes in kanamycin-resistant seedlings (R1 progeny, second generation) was measured and analyzed for each construct. The results are shown in Fig. 2a. This



**Fig. 2** **a** Expression analysis of constructs pKPG, pKPLG, pKPL7RG, pKPL7G by real-time qRT-PCR in stably transformed transgenic tobacco plants. Twelve independent lines were made for each construct. For transgenic plant assay, independent lines (R1 progeny, second generation) with Kan<sup>R</sup>:Kan<sup>S</sup> = 3:1 were selected. Samsun NN was used as the untransformed control. GUS activity (nmol 4-methylumbelliferone min<sup>-1</sup> mg<sup>-1</sup>; mean  $\pm$  SD) is presented for each construct. **b** Average initial copy number of GUS transcripts per 100 ng of total RNA in transgenic lines derived for the constructs as indicated, and untransformed Samsun NN (*Control*)

expression level was about 2-fold higher in pKPL7RG lines than in pKPG lines. Interestingly, GUS expression in plants carrying pKPL7RG was about 800-fold higher than that in plants carrying pKPL7G.

Real-time quantitative RT-PCR analysis was performed to determine the abundance of GUS transcripts in plants developed for the various chimeric constructs. The accumulation of GUS transcripts in pKPL7RG-plants was about 1.3-fold more than in pKPG-plants, about 35-fold higher than in pKPL7G-plants, and about 10-fold greater than in pKPLG-plants (Fig. 2b). Melting curves were analyzed for the sample (GUS) as well as the internal control ( $\alpha$ -tubulin) to test the specificity of the PCR product. The melting curve showed a smooth decline in fluorescence with an increase in temperature as the strands of double-stranded DNA dissociated and the bound SYBR Green was released. The melting temperature was seen as a single peak at 81 °C for GUS and 83 °C for  $\alpha$ -tubulin, indicating the homogeneity of the amplified products. These data indicate that p7R sequence increased the transcript level, and thus an increase in mRNA abundance is responsible for high expression of the GUS gene.

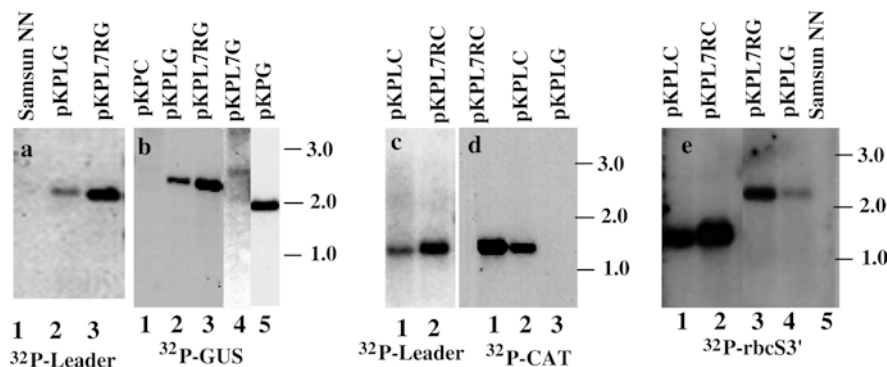
## Splicing of p7R-containing RNAs enhanced gene expression

To better understand the nature of the effects of p7R on reporter gene expression, transcripts arising from pKPG, pKPLG, pKPL7G, pKPL7RG, pKPLC, and pKPL7RC, were analyzed by northern blots. Plants containing the pKPLG construct showed an expected band of  $\approx$ 2.4 kb when probed with either the PCISV leader (Fig. 3a, lane 2), GUS (Fig. 3b, lane 2) or *rbcS3'* (Fig. 3e, lane 4). Plants carrying pKPL7RG had a transcript of  $\approx$ 2.3 kb that hybridized with the leader (Fig. 3a, lane 3), GUS (Fig. 3b, lane 3) or *rbcS3'* (Fig. 3e, lane 3) probes. This is about 500 nt smaller than the expected size. A small amount of a 2.8-kb RNA (the size expected for this gene) can be seen after a very long exposure (data not shown). This result suggests that as much as  $\approx$ 0.5 kb may be spliced out from the transcript encoded by pKPL7RG.

When RNA blots were probed with labeled p7, samples from pKPL7RG-plants did not show any signal (data not shown), indicating that the p7 sequence may be spliced out. Plants containing the construct pKPL7G that contain the p7 segment (sense strand of PCISV gene VII) gave a transcript of  $\approx$ 2.8 kb when probed with [<sup>32</sup>P]GUS (Fig. 3b, lane 4). Transgenic control plants containing pKPG gave a transcript of expected size  $\approx$ 2.1 kb (Fig. 3b, lane 5). The abundance of the 2.8-kb transcript was very low in pKPL7G-plants compared with that from pKPLG-plants, pKPL7RG-plants or pKPG-plants, (compare lane 4 with lanes 2, 3 and 5, respectively, in Fig. 3b). This suggests that the p7R sequence was spliced out as an intron from the pre-mRNA (corresponding to leader-p7R-GUS, size of  $\approx$ 2.8 kb).

Similar studies were performed with plants that carry CAT-containing constructs. When probed with the CAT sequence, plants transformed with pKPLC gave a transcript of expected size  $\approx$ 1.3 kb (Fig. 3d, lane 2). Plants

**Fig. 3a–e** Northern analysis of total RNA isolated from transgenic tobacco and untransformed tobacco (Samsun NN). Blots were hybridized with leader (a), GUS (b), leader (c), CAT (d), and *rbcS3'* (e) probes. The name of the construct used for transformation is at the *top* of each lane. Size (in kb) of the standard RNA marker (Millenium marker from Ambion) is on the *right-hand side* of each blot. Equal loading of RNA was confirmed by ethidium bromide staining of rRNA as indicator



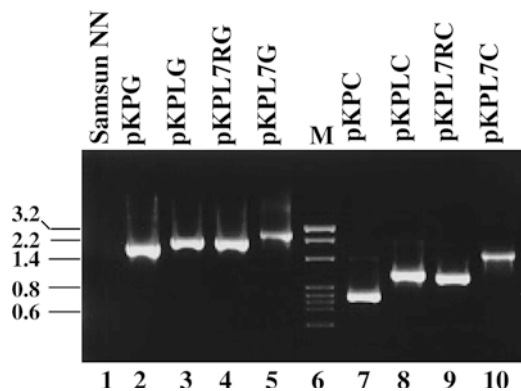
transformed with pKPL7RC showed a major transcript of size similar to that seen for pKPLC plants (Fig. 3d, lane 1); this is about 500 nt less than the predicted size ( $\approx 1.7$  kb). Plants carrying pKPLC and pKPL7RC gave a very similar size of transcripts when probed with either the leader sequence (Fig. 3c, lanes 1 and 2, respectively) or the rbcS3' sequence (Fig. 3e, lanes 1 and 2, respectively). Our results suggest that the p7R sequence is removed from the pre-mRNA. Splicing of p7R as an intron is also independent of reporter CAT and GUS gene sequences tested here. Therefore, the discrepancy between expected and actual mRNA size for pKPL7R-constructs does not depend on the nature of the reporter gene.

A northern blot of total RNA isolated from plants carrying pKPLC, pKPL7RC, pKPLG and pKPL7RG, when probed with rbcS3' sequences, showed the range of transcripts seen with other probes (Fig. 3e). Therefore, all of these mRNAs contain the rbcS3' polyA sequence. This argues against an alternative poly(A) site choice as an explanation for the discrepancies mentioned above.

As expected from the qRT-PCR analysis, the relative abundance of transcript from pKPL7RG-plants was severalfold higher than that from the pKPLG-plants (Fig. 3b, compare lane 3 with lane 2) and pKLP7G (Fig. 3b, compare lane 3 with lane 4). Similarly with the CAT reporter gene, pKPL7RC-plants showed greater transcript levels than pKPLC plants (Fig. 3c, compare lane 1 with lane 2). Taken together, our data indicate that the p7R sequence increased the abundance of the GUS and CAT transcripts. It is thus likely that an increase in mRNA abundance is responsible for high expression of these chimeric gene constructs.

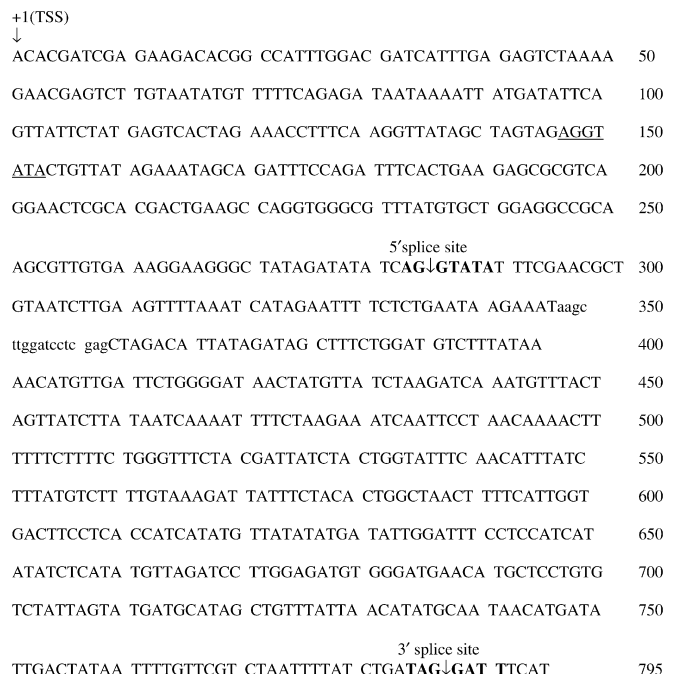
#### Identification of splicing sites through RT-PCR and DNA sequencing

In order to find the structural differences in various transcripts, RT-PCR analysis was conducted using



**Fig. 4** RT-PCR analysis of total RNA from tobacco plants developed for constructs as indicated. The general structure of these constructs is shown in Fig. 1. RT-PCR analysis is displayed for untransformed Samsun NN (lane 1), pKPG (lane 2), pKPLG (lane 3), pKPL7RG (lane 4), pKPL7G (lane 5), DNA size marker (M, lane 6; size in kb is indicated on the left), pKPC (lane 7), pKPLC (lane 8), pKPL7RC (lane 9) and pKPL7C (lane 10)

primers designed to amplify the full-sized mRNAs as described in Materials and methods. The pKPG-plants gave an  $\approx 1.8$ -kb GUS product of expected size (Fig. 4, lane 2). An RT-PCR product of size  $\approx 2.1$  kb (corresponding to leader with GUS; Fig. 4, lanes 3 and 4) was seen in samples from both pKPLG and pKPL7RG plant lines, as expected from the northern results. An  $\approx 2.5$ -kb band of expected size was obtained from the pKPL7G-plants. In addition, RT-PCR products of expected size  $\approx 0.7$  kb,  $\approx 1.0$  kb,  $\approx 0.95$  kb and  $\approx 1.4$  kb were obtained from plants carrying pKPC (lane 7), pKPLC (lane 8), pKPL7RC (lane 9) and pKPL7C (lane 10), respectively. DNA sequence analysis of this 2.1-kb RT-PCR fragment from pKPL7RG-plants and 0.95-kb RT-PCR fragment from pKPL7RC-plants showed that the entire 5'-end of the p7R sequence, except the terminal 8 nt from the 3'-end of p7R, and the 60 nt from the 3'-end of the leader sequence were absent (Fig. 5). The RT-PCR product from transcript of pKPLG-/ pKPLC-plants or pKPL7G-/ pKPL7C-plants lacked the deletion seen with p7R-derived constructs. The sequences of the RT-PCR DNA products from five independent transgenic lines harboring pKPL7RG or pKPL7RC constructs clearly showed that the p7R sequences were absent from the pre-mRNA. Invariably, the junction sequences at the 5'- and 3'-exon-intron border are CAG/GUAU and UAG/GAUU, respectively (Fig. 5). These fit the consensus splice site for plants, and strongly suggest that the missing sequence has been removed by splicing. The



**Fig. 5** DNA sequence of the PCISV full-length transcript-leader (coordinates 1 to 346), fused with the p7R sequence (coordinates 364 to 795) is presented from left to right in the 5' to 3' direction. The transcription start site (TSS, +1), leader-p7R junction sequence (lower case), 5' and 3' splicing sites (bold; cleavage point indicated by an arrow), and unused 5' splice site (underlined) are shown

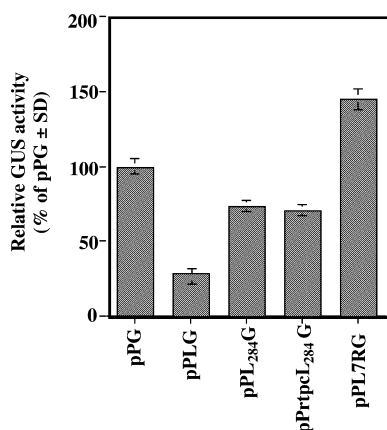
active 5' splice site (AG/GTATA) is at position (+283 to +289) from TSS. An identical consensus 5' splice site at position (+147 to +153) from TSS (Fig. 5) is not activated by the p7R sequence in this context. The consensus active 3' splice site (TAG/GATT) is located on the p7R sequence at position (+785 to +789) from the TSS (Fig. 5).

Intron-mediated enhancement (IME) is responsible for the higher expression in the p7R-derived constructs

In order to distinguish the role of splicing and the effect, if any, of the truncated leader (i.e. devoid of 3' terminal 60 nt that is spliced out from the leader), with and without the 3'-terminal 8 nt of p7R, constructs pPrtpcL<sub>284</sub>G and pPL<sub>284</sub>G were studied. The difference between these two constructs is that the latter is devoid of the 3'-terminal 8 nt of p7R. The pPrtpcL<sub>284</sub>G mRNA is identical in structure to the spliced mRNA from the construct pPL7RG. The levels of GUS expression obtained with constructs pPrtpcL<sub>284</sub>G and pPL<sub>284</sub>G were very similar (Fig. 6). Although the expression levels with both constructs were 2-fold higher than that obtained with the pPLG construct (which contains the entire leader sequence), they were about 50% less than that obtained with pPL7RG (Fig. 6). The level of GUS expression in pPL7RG was 42% higher than that obtained from pPG (where the GUS gene is under the strong P9 promoter). These results suggest that the higher level of expression of the downstream GUS gene in pPL7RG is an IME phenomenon.

#### Promoter specificity

In order to test the promoter specificity in splicing of p7R and subsequent enhancement of gene expression, the P9



**Fig. 6** Splicing of the intron is essential for maximum enhancement. The construct pPrtpcL<sub>284</sub>G contains the sequence as in the mature RNA after splicing. It is devoid of the 60-nt-long 3'-terminal leader sequence but contains the 3'-terminal 8 nt of the p7R sequence (Fig. 5). The construct pPL<sub>284</sub>G is devoid of the 60-nt-long 3'-terminal leader sequence. The construct pPLG contains the whole leader sequence. The PCISV strong P9 promoter directs the GUS gene in pPG. These constructs were generated as described in Materials and methods

promoter was substituted with other caulimovirus promoters and the resulting constructs were analyzed in a protoplast transient-expression assay. Results are shown in Table 1. The GUS activities of constructs with the respective promoters but without the PCISV leader and p7R sequence were taken as controls. The CaMV 35S, 19S promoters and the *Casava vein mosaic virus* (CVMV) FLt promoter (Verdaguer et al. 1996), as well as the *Mirabilis mosaic virus* (MMV) Sgt promoter (Dey and Maiti 2002), showed about 1.5- to 2-fold more activity than the respective controls, whereas the MMV FLt (Dey and Maiti 1999), *Figwort mosaic virus* (FMV) FLt (Maiti et al. 1997), and FMV Sgt (Bhattacharyya et al. 2002) promoters showed a slightly decreased expression level compared with controls. However, the size and sequence analysis of RT-PCR products derived from the total RNA isolated from electroporated protoplasts revealed that splicing occurred in those constructs containing any of the caulimovirus promoters tested in this context (data not presented).

## Discussion

p7R intron-mediated enhancement of expression is an IME phenomenon, and it is not gene dependent

Here we have documented that the antisense orientation of the PCISV ORF VII (denoted as p7R), in conjunction

**Table 1** Influence of various caulimovirus promoters on the p7R-intron-mediated enhancement of gene expression in tobacco (*Nicotiana tabacum* L. cv. Xanthi). Constructs were assayed in tobacco protoplast transient-expression experiments. The physical map of each GUS construct containing the caulimovirus promoter is shown: CaMV 35S promoter (35S), MMV FLt promoter (M12), FMV FLt promoter (F1), CVMV FLt promoter (C1), FMV Sgt promoter (FS3), MMV Sgt promoter (MS8), CaMV 19S promoter (19S), PCISV leader (L) and the antisense orientation of PCISV ORF VII (p7R). In these promoter constructs the 5'-EcoRI-P9-HindIII-3' promoter fragment in pPL7RG and pPG was replaced with other caulimovirus promoters as indicated. Each construct was assayed at least two times in two independent experiments. The average GUS activity is presented as fold enhancement of the respective control. The variation was within 10% of the presented value

General structure of promoter construct	Enhancement of GUS activity (-fold)
35S-GUS	1.0
35S-L-p7R-GUS	2.00
M12-GUS	1.0
M12-L-p7R-GUS	0.83
F1-GUS	1.0
F1-L-p7R-GUS	0.82
C1-GUS	1.0
C1-L-p7R-GUS	1.4
FS3-GUS	1.0
FS3-L-p7R-GUS	0.78
MS8-GUS	1.0
MS8-L-p7R-GUS	1.63
19S-GUS	1.0
19S-L-p7R-GUS	2.0

with the PCISV leader, can enhance the expression of a downstream gene by functioning as an intron. In this splicing process, the active 5' splicing donor site is located at +283 from the TSS on the leader sequence; and the 3' splice site is located on the p7R sequence at position (+785 to +791) from the TSS (Fig. 5). The p7R is spliced-out itself, along with the 3'-terminal 60 nt of the PCISV leader sequence.

The construct pPrtpcL<sub>284</sub>G yielded half of the expression obtained from pPL7RG (Fig. 6), even though both constructs should produce transcripts with the same sequence. The functional difference between the two constructs is that RNAs derived from the latter undergo splicing. Therefore, p7R-mediated enhancement of expression is an IME phenomenon. The IME phenomenon has been found to be more efficient in monocot plants rather than dicot plants (Tanaka et al. 1990; Vain et al. 1996) and it may be a gene-dependent process (Rethmeier et al. 1997). Our results clearly establish that the p7R sequence can function as an intron in dicot tobacco plants and may not be gene-dependent as it is able to enhance expression of both GUS and CAT.

It is believed that the small ORFs present in the 5' leader regions of the caulimovirus full-length transcript are inhibitory to the translation process of downstream genes. We find that the deletion of the 3'-terminal 60 nt from the PCISV leader (construct pPL<sub>284</sub>G) can enhance expression of a downstream gene at least 2-fold compared to pPLG (Fig. 6). The four small ORFs present in the leader sequence are not affected by the deletion of 60 nt from its 3'-end terminal. It may be concluded that small ORFs on the PCISV leader are not solely responsible for inhibiting the translation of the downstream genes. Rather, the 3'-terminal 60 nt of the PCISV leader may have some important role in secondary structure formation that controls the translation of downstream genes.

The GUS expression level in construct pPrtpcL<sub>284</sub>G (general structure: P9-leader1-284 + 3'-terminal p7R 8 nt-GUS) did not differ from that of the construct pPL<sub>284</sub>G (general structure: P9-leader1-284-GUS; Fig. 6). The latter construct has only the 3'-truncated leader and is devoid of the 3'-terminal 8 nt of p7R. This suggests that the 8-nt terminal sequence of p7R has no appreciable effect on the expression of a gene in this context.

We have also shown (Table 1) that the splicing of p7R occurs independently of the promoter used in this context; however, the splicing event does not always lead to an enhancement of gene expression. Similar results have been reported for the Adh1-S intron in maize cells (Luehrsen and Walbot 1991). More work will be needed to evaluate the mechanism of the IME phenomenon.

Uracil (U)-richness of the p7R intron may play a role in activation of splicing sites

There are ambiguities about the mechanism of IME (Rethmeier et al. 1997; Rose and Beliakoff 2000). In

plants it is a common feature that introns act post-transcriptionally to increase mRNA accumulation. It has been suggested that a highly U-rich intron stabilizes the nascent transcript by providing targets for protective RNA-binding factors (Gniadkowski et al. 1996), or heterogeneous nuclear ribonucleoproteins and spliceosomal assembly (Krecic and Swanson 1999). The U-rich p7R (U content 42.1%) may play the same role in promoting the accumulation of higher amounts of processed mRNA.

Plant introns are very AU-rich. Although p7R is not of plant origin, it functions as an intron in plants probably due to its AU-richness (70.6%). The constructs pKPL7G and pKPL7RG have the same p7 regions, and the same leader sequence; moreover, they have the same junction sequence between the leader and p7 or p7R, as well as between p7 or p7R and the reporter gene. The difference between the two constructs involves only the sense and antisense strand of PCISV ORF VII (p7). The sense strand of ORF VII (p7) did not activate the 5' splicing donor site in the PCISV leader. However, p7R activated the 5' splicing donor site in the leader so efficiently that the level of unspliced pre-mRNA was very low, maybe less than 1%. The AU-contents for p7 and p7R are the same (70.6%) whereas they differ in U-content (28.5% in p7 and 42.1% in p7R). It may be that the intron nature of p7R in plants is due to the U-richness and/or number of U-rich tracts in the sequence. Sequence comparison of different plant introns also showed that U-richness of the intron sequence, but not the AU-richness, is the defining feature of a plant intron (Ko et al. 1998), especially in maize. U-rich motifs play a critical role in the recognition of the 3' splicing acceptor site in *Saccharomyces cerevisiae* (Guthrie 1991; Umen and Guthrie 1995) and in plants (Bayton et al. 1996).

Surrounding junction sequence of two identical donor sites in the PCISV leader

The PCISV leader sequence has two identical 5' splice junction sequences (AGGUAUA) at +148 and +284 nt from the TSS (Fig. 5). However, p7R can only activate the 5' splice site at the +284 nt position. The other donor site (+148 from TSS) was not activated in this context. An identical UAUAG motif at the -14 to -9 nt upstream position of both donor sequences AG/GUAUA (i.e. +148 nt and +284 nt position of leader from TSS) is present. The region upstream of the UAUAG motif (-45 to -16) is highly purine-rich (73.4%) in the functional 5' donor site (+284 nt position of leader) whereas the purine content is 46.7% upstream of the other (unused) 5' donor site (+148 nt position of leader). This highly purine-rich sequence upstream of the splicing donor site might have some interacting role with U-rich intronic sequences or U-rich-binding transacting factors. It had been established that purine-rich motifs located upstream and downstream of the exon-intron border regulate alternative splice-site selection in plants

(Egoavil et al. 1997), chicken troponin T (TNT) exon 16 (Wang et al. 1995), mouse immunoglobulin exonM2 (Ramchatesingh et al. 1995), and avian ASLV env exon (Tanaka et al. 1994). However, such a purine-rich sequence is absent downstream of the splicing donor-site of p7R.

The role of a branch-point sequence (YUNAN) in plant-intron splicing is poorly understood, as it is relatively less conserved (Brown et al. 1996) than that in animal and yeast systems. A number of such conserved motifs are present on the entire length of the p7R sequence. The p7R has a branch-point sequence in the region -30 to -50 nt from the 3' splicing site.

Recently, a number of reports on splicing events in plants have addressed the activation of 5' and 3' splicing sites (McCullough et al. 1993; Watakabe et al. 1993; Bayton et al. 1996; Simpson and Filipowicz 1996; Egoavil et al. 1997; McCullough and Schuler 1997; Bourgeois et al. 1999; Alvarez and Wise 2001), as well as the plant intron recognition signal (Luehrsen and Walbot 1994a, 1994b; Ko et al. 1998; Lambermon et al. 2000; Romfo et al. 2000). However, little is known about the nature of viral sequences that function as introns or exons in the plant nucleus. The viral sequences, reported here, may serve as a model for identifying the *cis* and *trans* elements required for plant pre-mRNA splicing.

It is evident that the p7R in conjunction with part of the PCISV leader sequence can function as a model intron in plants. The PCISV leader with the p7R sequence can be used to construct genetic tools for intron-mediated enhanced expression of genes in plants. p7R is not a natural plant intron but it can be used as a model system for plant intron studies. It will be interesting to identify the *trans*-acting factors responsible for plant pre-mRNA splicing as well as activation of the 5' splice site.

**Acknowledgments** We are greatly indebted to the Kentucky Tobacco Research and Development Center for facilities and support. The KTRDC Project 5-41132 to IBM supported this research. We thank Drs. Arthur Hunt, Thomas Pirone, H. Maelor Davies and the anonymous reviewer for critical reading and valuable comments for improving the manuscript.

## References

- Alvarez CJ, Wise JA (2001) Activation of a cryptic 5' splice site by U1 snRNA. *RNA* 7:342-350
- Bayton CE, Potthoff SJ, McCullough AJ, Schuler MA (1996) U-rich tracts enhance 3' splice recognition in plant nuclei. *Plant J* 10:703-711
- Bhattacharyya S, Dey N, Maiti IB (2002) Analysis of *cis*-sequence of subgenomic transcript promoter from the *Figwort mosaic virus* and comparison of promoter activity with the *Cauliflower mosaic virus* promoters in monocot and dicot cells. *Virus Res* 90:47-62
- Bourgeois CF, Popielarz M, Hildwein G, Stevenin J (1999) Identification of a bidirectional splicing enhancer: differential involvement of SR proteins in 5' or 3' splice site activation. *Mol Cell Biol* 19:7347-7356
- Bradford MM (1976) A rapid and sensitive method for quantification of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72:248-254
- Brown JWS, Smith P, Simpson CG (1996) Arabidopsis consensus intron sequences. *Plant Mol Biol* 32:531-535
- Buchman AR, Berg P (1988) Comparison of intron-dependent and intron-independent gene expression. *Mol Cell Biol* 8: 4395-4405
- Callis J, Fromm M, Walbot V (1987) Introns increase gene expression in cultured maize cells. *Genes Dev* 1:1183-1200
- Christensen AH, Quail PH (1996) Ubiquitin promoter-based vectors for high-level expression of selectable and/or screenable marker genes in monocotyledonous plants. *Transgenic Res* 5:213-218
- Christensen AH, Sharrock RA, Quail PH (1992) Maize polyubiquitin genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation. *Plant Mol Biol* 18:675-689
- Chung S, Perry R (1989) Importance of introns for expression of mouse ribosomal protein gene rpL32. *Mol Cell Biol* 9:2075-2082
- Clancy M, Vasil V, Hannah LC, Vasil IK (1994) Maize *Shrunken-1* intron and exon regions increase gene expression in maize protoplasts. *Plant Sci* 98:151-161
- Dean C, Favreau M, Bond-Nutter D, Bedbrook J, Dunsmuir P (1989) Sequences downstream of translation start regulate quantitative expression of two petunia *rbcS* genes. *Plant Cell* 1:201-208
- Dennis ES, Gerlach WL, Pryor AJ, Bennetzen JL, Inglis A, Llewellyn D, Sachs MM, Ferl RJ, Peacock WJ (1984) Molecular analysis of the alcohol dehydrogenase (*Adh I*) gene of maize. *Nucleic Acids Res* 12:3983-4000
- Dey N, Maiti IB (1999) Structure and promoter/leader deletion analysis of mirabilis mosaic virus (MMV) full-length transcript promoter in transgenic plants. *Plant Mol Biol* 40:771-782
- Dey N, Maiti IB (2002) Promoter deletion and comparative expression analysis of the *Mirabilis mosaic caulimovirus* (MMV) sub-genomic transcript (Sgt) promoter in transgenic plants. *Transgenics* 4:35-53
- Donath M, Mendel R, Cerff R, Martin W (1995) Intron-dependent transient expression of the maize *GapA1* gene. *Plant Mol Biol* 28:667-676
- Ducasse DA, Shepherd RJ (1995) Systemic infection of solanaceous hosts by peanut chlorotic streak caulimovirus is temperature dependent and can be complemented by coinfection with figwort mosaic caulimovirus. *Phytopathology* 85:286-291
- Dugdale B, Becker DK, Harding RM, Dale JL (2001) Intron-mediated enhancement of the banana bunchy top virus DNA-6 promoter in banana (*Musa* spp.) embryogenic cells and plants. *Plant Cell Rep* 20:220-226
- Egoavil C, Marton HA, Baynton CE, McCullough AJ, Schuler MA (1997) Structural analysis of elements contributing to 5' splice site selection in plant pre-mRNA transcripts. *Plant J* 12:971-980
- Futterer J, Hohn T (1991) Translation of a polycistronic mRNA in the presence of the cauliflower mosaic virus transactivator protein. *EMBO J* 10:3887-3896
- Gniadkowski M, Hemmings-Mieschak M, Klahare U, Liu H-X, Filipowicz W (1996) Characterization of intronic uridine-rich sequence elements acting as possible targets for nuclear proteins during pre-mRNA splicing in *Nicotiana plumbaginifolia*. *Nucleic Acids Res* 24:619-627
- Gorman CM, Moffat LF, Howard BH (1982) Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. *Mol Cell Biol* 2:1044-1051
- Guthrie C (1991) Messenger RNA splicing in yeast: clues to why the spliceosome is a ribonucleoprotein. *Science* 253:157-163
- Huang Z, Fasco MJ, Kaminsky LS (1996) Optimization of DNaseI removal of contaminating DNA from RNA for use in quantitative RNA-PCR. *BioTechniques* 20:1012-1020

- Jefferson RA, Kavanagh TA, Bevan MW (1987) GUS fusion:  $\beta$ -glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J* 6:3901–3907
- Ko CH, Brendel V, Taylor RD, Walbot V (1998) U-richness is a defining feature of plant introns and may function as an intron recognition signal in maize. *Plant Mol Biol* 36:573–583
- Koziel MG, Carozzi NB, Desai N (1996) Optimizing expression of transgenes with an emphasis on post-transcriptional events. *Plant Mol Biol* 32:393–405
- Krecic AM, Swanson MS (1999) hnRNP complexes: composition, structure and function. *Curr Opin Cell Biol* 11:362–371
- Lambermon LHM, Simpson G, Wieczorek Kirk DA, Hemmings-Mieszczyk M, Klahre U, Filipowicz W (2000) UBP1, a novel hnRNP-like protein that functions at multiple steps of higher plant nuclear pre-mRNA maturation. *EMBO J* 7:1638–1649
- Luehrsen KR, Walbot V (1991) Intron enhancement of gene expression and the splicing efficiency of introns in maize cells. *Mol Gen Genet* 225:81–93
- Luehrsen KR, Walbot V (1994a) Addition of A- and U-rich sequence increases the splicing efficiency of a deleted form of a maize intron. *Plant Mol Biol* 24:449–463
- Luehrsen KR, Walbot V (1994b) Intron creation and polyadenylation in maize are directed by AU-rich RNA. *Genes Dev* 8:1117–1130
- Maas C, Laufs J, Grant S, Korfhage C, Werr W (1991) The combination of a novel stimulatory element in the first exon of the maize *shrunken-1* gene with the following intron enhances reporter gene expression 1000-fold. *Plant Mol Biol* 16:199–207
- Maiti IB, Shepherd RJ (1998) Isolation and expression analysis of peanut chlorotic streak caulimovirus (PCISV) full-length transcript (FLt) promoter in transgenic plants. *Biochem Biophys Res Commun* 224:440–444
- Maiti IB, Murphy JF, Shaw JG, Hunt AG (1993) Plants that express a potyvirus proteinase are resistant to virus infection. *Proc Natl Acad Sci USA* 90:6110–6114
- Maiti IB, Gowda S, Kierman J, Ghosh SK, Shepherd RJ (1997) Promoter/leader analysis and plant expression vectors with the figwort mosaic virus (FMV) full length transcript (FLt) promoter containing single and double enhancer domains. *Transgenic Res* 6:143–156
- Maiti IB, Richins RD, Shepherd RJ (1998) Gene expression regulated by gene VI of caulimovirus: transactivation of downstream genes of transcript by gene VI of peanut chlorotic streak virus in transgenic tobacco. *Virus Res* 57:113–114
- Maiti IB, Dey N, Pattanaik S, Dahlman DL, Rana RL, Webb BA (2003) Antibiosis-type insect resistance in transgenic plants expressing a teratocyte secretory protein (TSP14) gene from a hymenopteran endoparasite (*Microplitis croceipes*). *Plant Biotechnol J* (in press)
- Mascarenhas D, Mettler IJ, Pierce DA, Lowe HW (1990) Intron-mediated enhancement of heterologous gene expression in maize. *Plant Mol Biol* 15:913–920
- McCullough AJ, Schuler MA (1997) Intronic and exonic sequences modulate 5' splice site selection in plant nuclei. *Nucleic Acids Res* 25:1071–1077
- McCullough AJ, Lou H, Schuler MA (1993) Factors affecting authentic 5' splice site selection in plant nuclei. *Mol Cell Biol* 13:1323–1331
- McElroy D, Zhang W, Cao J, Wu R (1990) Isolation of an efficient actin promoter for use in rice transformation. *Plant Cell* 2:163–171
- Meredith J, Storti RV (1993) Developmental regulation of the *Drosophila* tropomyosin II gene in different muscles is controlled by muscle-type-specific intron enhancer elements and distal and proximal promoter control elements. *Dev Biol* 159:500–512
- Okkema PG, Harrison SW, Plunger V, Aryana A, Fire A (1993) Sequence requirements for myosin gene expression and regulation in *Caenorhabditis elegans*. *Genetics* 135:385–404
- Ramchatesingh J, Zahler AM, Neugebauer KM, Roth MB, Cooper TA (1995) A subset of SR proteins activates splicing of the cardiac troponin T alternative exon by direct interactions with an exonic enhancer. *Mol Cell Biol* 15:4898–4907
- Rethmeier N, Seurinck J, Van Mantagu M, Corelissen M (1997) Intron-mediated enhancement of transgene expression in maize is a nuclear gene dependent process. *Plant J* 12:896–899
- Richins RD (1993) Organization and expression of the peanut chlorotic streak virus genome. Dissertation, University of Kentucky, Lexington, p 153
- Ririe KM, Rasmussen RP, Wittwer CT (1997) Product differentiation by analysis of DNA melting curves during the polymerase chain reaction. *Anal Biochem* 245:154–160
- Romfo CM, Alvarez CJ, Van Heeckeren WJ, Webb CJ, Wise JA (2000) Evidence for splice site pairing via intron definition in *Schizosaccharomyces pombe*. *Mol Cell Biol* 20:7955–7970
- Rose AB, Beliakoff JA (2000) Intron-mediated enhancement of gene expression independent of unique intron sequence and splicing. *Plant Physiol* 122:535–542
- Rose AB, Last RL (1997) Introns act post-transcriptionally to increase expression of the *Arabidopsis thaliana* tryptophan pathway gene *PAT1*. *Plant J* 11:455–464
- Rothnie HM, Chapdelaine Y, Hohn T (1994) Pararetrovirus and retrovirus: a comparative review of viral structure and gene expression strategy. *Adv Virus Res* 44:1–67
- Scharl CL, Byrd AD, Benzion G, Altschuler MA, Hilderbrand DF, Hunt AG (1987) Design and construction of a versatile system for the expression of foreign genes in plants. *Gene* 61:1–11
- Shepherd RJ (1989) Biochemistry of DNA plant viruses. In: Marcus A (ed) *The biochemistry of plants, A comprehensive treatise*, vol 15. Academic Press, New York, pp 563–616
- Simpson GG, Filipowicz W (1996) Splicing of precursors to mRNA in higher plants: mechanism, regulation and subcellular organization of the spliceosomal machinery. *Plant Mol Biol* 32:1–41
- Sinibaldi RM, Mettler IJ (1992) Intron splicing and intron-mediated enhanced expression in monocots. In: Cohn WE, Moldave K (eds) *Progress in nucleic acid research and molecular biology*, vol 42. Academic Press, New York, pp 229–257
- Snowden KC, Buchholz WG, Hall TC (1996) Intron position affects expression from the *tpi* promoter in rice. *Plant Mol Biol* 31:689–692
- Tanaka A, Mita S, Ohta S, Kyozuka J, Shimamoto K, Nakamura K (1990) Enhancement of foreign gene expression by a dicot intron in rice but not in tobacco is correlated with an increased level of mRNA and an efficient splicing of the intron. *Nucleic Acids Res* 18:6767–6770
- Tanaka K, Watakabe A, Shimura Y (1994) Polypurine sequences within a downstream exon function as a splicing enhancer. *Mol Cell Biol* 14:1347–1354
- Umen JG, Guthrie C (1995) A novel role for a U5 snRNP in 3' splice site selection. *Genes Dev* 9:855–868
- Vain P, Finer KR, Enggler DE, Pratt RC, Finer JJ (1996) Intron-mediated enhancement of gene expression in maize (*Zea mays* L.) and bluegrass (*Poa pratensis* L.). *Plant Cell Rep* 15:189–194
- Verdaguer B, de Kochko A, Beachy RN, Fauquet C (1996) Isolation and expression in transgenic tobacco and rice plants of the vein mosaic virus (CVMV) promoter. *Plant Mol Biol* 31:1129–1139
- Wang Z, Hoffmann HM, Grabowski PJ (1995) Intrinsic U2AF binding is modulated by exon enhancer signals in parallel with changes in splicing activity. *RNA* 1:21–35
- Watakabe A, Tanaka K, Shimura Y (1993) The role of exon sequences in splice site selection. *Genes Dev* 7:407–418