

In vivo analysis of plant pre-mRNA splicing using an autonomously replicating vector

Andrew J. McCullough, Hua Lou, and Mary A. Schuler

Department of Plant Biology, University of Illinois, Urbana, IL 61801, USA

Received February 10, 1991; Revised and Accepted April 15, 1991

ABSTRACT

In this paper, we demonstrate that an autonomously replicating plant expression vector can be used for analysis of pre-mRNA splicing determinants in intact dicot cells. This vector system relies on the *Agrobacterium*-mediated transfection of leaf discs with the A component of the geminivirus tomato golden mosaic virus (TGMV). Insertion of intron sequences between viral promoter and terminator sequences results in the production of high levels of pre-mRNA transcripts that are effectively and accurately spliced *in vivo*. Introns from the soybean B-conglycinin gene are spliced at >95% efficiency indicating that the high expression levels of precursor RNA do not exceed the intron splicing capacity of these cells. Introns from the pea and wheat rbcS genes are spliced at 85% and 73% efficiency, respectively, indicating that tobacco leaf disc nuclei are capable of effectively and accurately processing particular dicot and monocot introns. Inclusion of a dicot intron in an engineered construct results in a five-fold enhancement of the level of mRNA stably expressed in dicot nuclei.

INTRODUCTION

In eukaryotes, nuclear genes are interrupted by introns that must be accurately excised from precursor transcripts to produce mature, functional mRNA products. Although all higher organisms contain genes with introns, the mechanisms of intron excision and the sequences which specify the process have been defined in only a small number of vertebrate and yeast cell types. Sequence comparisons in a number of metazoa have indicated that three sequence elements are conserved for the molecular ligation of exons. In mammalian systems, the first of these elements, the 5' intron boundary, consists of a relatively conserved set of nine nucleotides invariably containing a GT dinucleotide at the exon/intron boundary (review: 1). The second element critical for intron splicing in metazoa is a conserved branch sequence typically located between 20 and 50 nucleotides upstream from the 3' intron/exon boundary (1, 2). The third element, present at the 3' end of the intron, is an AG dinucleotide preceded by an extended polypyrimidine tract (1). Although the degree of sequence conservation in these elements varies in metazoan and yeast introns, *in vivo* and *in vitro* experiments have implicated all three of these conserved elements in the process of intron excision (3).

The complex intron splicing process is carried out by a large protein-RNA complex in which the precursor RNA is associated with small, nuclear ribonucleoprotein particles (snRNPs). Three snRNP particles, containing the U1, U2 and U5 small nuclear RNAs (snRNAs), provide the fundamental recognition factors required for the intron excision events (3). The U1 snRNP interacts directly with the 5' splice site (4–6) and the U2 snRNP interacts with the conserved branch sequence within the intron (5, 7). The U5 snRNP particle has been implicated in recognition of the 3' splice junction via its associated snRNP proteins (8). Two additional snRNAs, associated in the U4/U6 snRNP, participate in the splicing process (9, 10).

Although it might be logical to assume that, at some molecular level, all eukaryotes share a common mechanism for pre-mRNA splicing, a variety of *in vitro* and *in vivo* gene transfer experiments indicate that critical differences exist between mammalian, yeast and plant RNA processing events. This is particularly evident in *in vitro* experiments which have demonstrated that, although a few plant introns can be excised in mammalian nuclear extracts (11–14), many plant introns are not processed accurately in mammalian extracts (13, 14). The individual selectivities of the RNA processing systems extend to monocotyledonous and dicotyledonous plants and have prevented the efficient *in vivo* expression of some monocot genes containing introns in transgenic dicot plants. Introduction of a monocot gene encoding the small subunit of ribulose 1, 5-bisphosphate carboxylase of wheat or a fragment from the alcohol dehydrogenase gene of maize into transgenic tobacco plants resulted in the accumulation of intron-containing pre-mRNAs (15). Introduction of dicot introns, such as the small subunit of ribulose bisphosphate carboxylase intron of pea, into transgenic dicot plants resulted in the efficient and accurate processing of the precursor RNA into mature mRNA (15).

These results have suggested that *in vivo* the intron splicing machineries are interchangeable in a wide variety of dicots but not between monocots and dicots. Until now, only two systems have been available for further analysis of the precise sequences required for plant intron splicing. The first system, utilized by Keith and Chua (15), involves the introduction of intron-containing gene constructs into stably transformed plants and the subsequent long-term analysis of transcripts produced from single-copy genes integrated into the nuclear genome. The accurate quantitation of splicing in this system is complicated by variable expression levels which result from positional variation in the chromosomal integration sites. The second system involves the

introduction of intron-containing gene constructs into tobacco or maize protoplasts by polyethylene glycol transfection (16, 17) or *O. violaceus*, *N. tabacum*, *N. plumbaginifolia* or *Z. mays* protoplasts by electroporation (17, 18). In the latter system, precursor RNAs, which are transcribed from nonreplicating vectors transfected into the protoplasts, can be assayed more readily for splicing. But, because these vectors are only transiently expressed in protoplasts, the splicing efficiencies defined in this system reflect the integrity of the splicing apparatus in plant cells subjected to substantial stress.

In this paper, we describe another system that can be used for efficient analysis of intron splicing determinants in intact dicot leaf cells. This system, developed by Hanley-Bowdoin et al. (19) for transient expression of heterologous genes, relies on the *Agrobacterium*-mediated transfection of leaf discs with the A component of the geminivirus tomato golden mosaic virus (TGMV). After introduction into plant cells, the A component replicates autonomously in nuclei (20, 21) amplifying the number of templates available for transcription and, as a result, increasing the abundance of transcripts available for analysis. The levels of transcript produced from the geminivirus expression vector consistently parallel the accumulation of TGMV DNA and can be readily detected in intact plant cells between two and fourteen days after transfection (19). In our estimation, this type of vector is eminently suitable for the analysis of the sequences required for plant intron splicing because engineered transcripts can be generated in high abundance in the nuclei of intact plant cells. To test the feasibility of the system and produce high levels of intron-containing precursor RNAs, we have inserted full-length exon/intron/exon sequences between the constitutively-expressed TGMV coat protein promoter and terminator sequences. The resulting fusion transcripts resemble other RNA polymerase II transcripts in that they contain 5' and 3' nontranslated sequences and polyadenylation signals derived from the coat protein gene. These features presumably minimize the differences between endogenous and engineered precursor transcripts produced in the plant nuclei and allow splicing efficiencies to be defined in reasonably natural nuclear environments. In this paper, we demonstrate the utility of this vector for intron splicing analysis. In addition, we demonstrate that the intron splicing efficiencies defined in this system reflect sequences within particular introns rather than the overall efficiency of the RNA processing machinery.

MATERIALS AND METHODS

Cloning strategies

The pMON458 vector used in these studies is a derivative of the pMON351 vector (22), in which a BglII-AsuII restriction fragment containing the AR1 coat protein coding sequences has been replaced by a synthetic linker containing multiple restriction sites. The sequences in pMON458 used for our insertion of intron sequences are shown in Fig. 1C.

To subclone the pea (*Pisum sativum* L.) rbcS3A-1 into pMON458, the rbcS3A gene (23) was PCR amplified using the 3A-1 5' and 3A-1 3' primers which share 18 nt. of homology with the ends of exons 1 and 2, respectively. Each of the 5' primers used for subcloning contains a BamHI adapter sequence and the 3' primer contain nested BglII and EcoRI adapter sequences. The PCR amplification product, extending from the first nucleotide of exon 1 to the last nucleotide of exon 2, was

subcloned into the unique BglII site of pMON458 (Fig. 1) to generate pMON458-rbcS3A-1. Using similar PCR amplification strategies, the entire coding sequence of the wheat (*Triticum aestivum* L.) rbcSw4.3 gene (24) was subcloned into pMON458, generating pMON458-rbcS4.3. The soybean (*Glycine max* L.) β -conglycinin gene (25) was subcloned from a site within intron 2 to the last nucleotide of exon 6 generating pMON458- β con85-132.

Agrobacterium and leaf disc transfections

pMON458 constructs were mobilized into the *Agrobacterium tumefaciens* strain GV3111SE by the triparental mating scheme (26).

Nicotiana benthamiana leaves were collected from pre-flowering plants, cut into discs and plated on MS104 media [MS salts supplemented with B5 vitamins, 3% sucrose, 1 μ g/ml NAA,

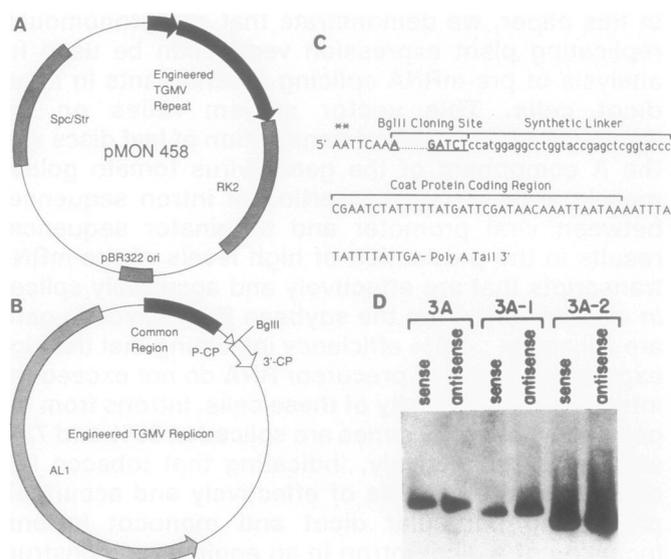


Figure 1. TGMV-based transient expression vector. (A) Diagrammatic representation of the pMON458 vector. The pMON458 vector, described by Gardiner et al. (22), contains one and one-half copies of an engineered TGMV A component (solid arrows) inserted into the binary *Agrobacterium* vector pMON505. The pertinent features of the pMON vector, shown with stippled boxes, include a pBR322 origin of replication for *E. coli*, an RK2 origin of replication for *Agrobacterium*, and streptomycin/spectinomycin (Str/Spc) resistance for selection in *E. coli* and *Agrobacterium*. (B) Representation of the engineered TGMV replicon in plant nuclei. The engineered TGMV replicon derived by homologous recombination of the pMON458 vector in *Nicotiana benthamiana* nuclei is shown with open boxes designating the AR1 coat protein promoter (P CP) and terminator (3' CP) sequences. Constructs for expression are ligated into the unique BglII site downstream of the AR1 viral coat protein promoter. The stippled box represents the AL1 transcription unit which codes for a protein required for viral DNA replication (21); the solid box represents the TGMV common region. (C) Sequence of engineered transcripts. Transcripts for intron analysis derived from BamHI:BglII insertion of heterologous sequences into the engineered TGMV coat protein gene (panel B) contain 8 nucleotides derived from the 5' nontranslated region of the TGMV coat protein gene (AR1), 31 nucleotides of linker DNA, 33 nucleotides of TGMV coat protein coding sequence, and the 3' nontranslated region extending to the putative polyadenylation site. The unique BglII restriction site used for our cloning is shown in underlined bold letters. The asterisks designate the coat protein transcript initiation sites defined in Sunter et al. (44). (D) Southern analysis of TGMV replicons isolated from transfected leaf discs. pMON458 constructs containing intron 1 (3A-1), intron 2 (3A-2), or both introns (3A) of the pea rbcS3A gene oriented in the sense or antisense direction were transfected into *N. benthamiana* leaf discs as described in Materials and Methods. Total DNA was isolated five days after transfection, digested with BglII to linearize the TGMV replicons, electrophoresed on a 1% agarose gel, transferred to Genescreen, and hybridized with a 32 P-labeled TGMV AL1 gene probe.

0.1 µg/ml BA]. The leaf discs were cultured on MS104 media for 2 days and subsequently infected with the appropriate *Agrobacterium tumefaciens* strains using standard methods (26). *Agrobacterium*-infected leaf discs were maintained on MS104 plates under lights for 1 to 9 days post-transfection, frozen in liquid nitrogen and stored at -80°C prior to nucleic acid isolation.

Nucleic acid preparation

Tissue from 10–12 leaf discs was ground for 1 minute using a Mini-Beadbeater (Biospec Products) in two volumes of buffer containing equal volumes of phenol and [100 mM LiCl, 100 mM Tris-HCl (pH 8.0), 10 mM EDTA, 1% SDS]. Samples were centrifuged for 5 minutes at 12,000×g to separate the aqueous and organic phases and the aqueous phase was re-extracted with an equal volume of phenol-chloroform (1:1). High molecular weight RNAs were precipitated with 2M LiCl on wet ice for 8–12 hr. Residual DNA was removed by digestion with 5U RNase-free DNAase for 60 min at 37°C.

Polymerase chain reaction (PCR) analysis

For analysis of RNA transcripts, first strand cDNA synthesis and PCR amplification were done in a single reaction mixture containing 1 µg total RNA, 5mM KCl, 10 mM Tris-HCl (pH 8.4), 2.5 mM MgCl₂, 20 µg/ml gelatin, 200 µM each dNTP, 2 units AMV Reverse Transcriptase, 2 units Taq DNA polymerase, 40 units RNasin and 100 pmoles of each appropriate oligonucleotide primer. The first strand cDNA template was synthesized for 30 minutes at 50°C and was subsequently amplified by 25 cycles of PCR. Each PCR cycle consisted of a 94°C denaturation for 1 minute, 60°C annealing for 2 minutes and 72°C extension for 2 minutes. PCR products were electrophoresed in 2% agarose gels containing 1×TBE buffer and visualized by ethidium bromide staining.

DNA sequence analysis

Using the BamHI and NotI sites in the adapter sequences, the PCR amplified products corresponding to the spliced transcripts were cloned into the Bluescript II SK+ vector (Stratagene) and sequenced using T7 DNA polymerase (Sequenase, U.S. Biochemicals) and the 3A-exon 2 and wheat-exon 1 sequencing primers (Table 1) following protocols supplied by the manufacturer.

Ribonuclease protection analysis

Antisense ³²P UTP-labeled RNA probes (Fig. 2) were analyzed using a RNase Protection Analysis kit (Ambion Inc.). 5–10 µg of total RNA were co-precipitated with 90,000 cpm of ³²P-labeled antisense RNA probe. The protected hybrids were fractionated on 6% acrylamide:bisacrylamide (38:2), 8.3 M urea denaturing gels and visualized by autoradiography. For quantification, radioactivity in excised bands was determined by liquid scintillation spectroscopy. The splicing efficiencies for each transcript were defined after adjusting the background-corrected spliced and unspliced cpm for differences in UMP content.

RESULTS

TGMV-based transient expression vector

The pMON458 vector used throughout this study contains one and one-half copies of an engineered tomato golden mosaic virus (TGMV) A component inserted into the pMON505 binary vector (19, 22). The wildtype geminivirus TGMV A component contains AL1, AL2, AL3 and coat protein (AR1) genes (27). In the engineered pMON458 vector (Fig. 1), the entire coat protein coding sequence except for 33 nucleotides at the carboxy-terminus has been deleted to allow for insertion of heterologous sequences between the AR1 promoter and terminator sequences (19, 22). As shown in Figure 1C, AR1 transcripts from the recombinant pMON458 vectors are composed of the 5' nontranslated region of AR1, inserted exon and intron sequences, synthetic linker sequences, 33 coding nucleotides at the carboxy-terminus of AR1 and the AR1 3' nontranslated region extending to the polyadenylation site (Fig. 1C, J.S. Elmer, communication).

Recombinant pMON458 vectors are introduced into *N. benthamiana* nuclei by *Agrobacterium*-mediated transformation. Homologous recombination between the repeated TGMV sequences in pMON458 releases one complete copy of the engineered TGMV A component containing the heterologous intron sequences. Because the engineered TGMV A component produces the functional AL1 gene product required for viral replication (21) and contains an origin of replication, it is capable of autonomous replication to approximately 100,000 copies per cell in *Nicotiana benthamiana* (S. Jobling, communication). The high copy number of the TGMV-based vector coupled with efficient transcription from the coat protein promoter results in the production of large amounts of AR1 fusion transcripts in transfected cells.

Table 1. PCR and sequencing oligonucleotides

Designation	Nucleotide Sequence (5'-3')	Complementarity
3A1-5'	cacagatgccATGGCTTCTATGATATCC	5' end of pea rbcS3A exon 1
3A1-3'	ccccgaattcagatctCTCCAACCTCAAATTCCAA	3' end of pea rbcS3A exon 2
wheat-5'	cacagatgccATGGCCCCGCCGTGATGGC	5' end of wheat rbcSw4.3 exon 1
wheat-3'	ccccgaattcagatctTAGGCCTTGCCGGACTCCT	3' end of wheat rbcSw4.3 exon 2
85-5'	cacagatgccTTTCCAAATCTTAAGATCAA	soybean B-con intron 2
132-5'	cacagatgccGGAGCTCTTTTCTACCACA	5' end of soybean B-con exon 5
132-3'	ccccgaattcagatctCAGTAAAAAGCCCTCAAAA	3' end of soybean B-con exon 6
coat-not	ccccgcccgcGAATCATAAAAAATAGATTCG	3' end of TGMV coat protein
3A-exon 2	CAAATAGGAAAAGAGTCTC	pea rbcS3A exon 2
wheat exon 1	AGCGTCAGCAAAGCGGGA	wheat rbcSw4.3 exon 1

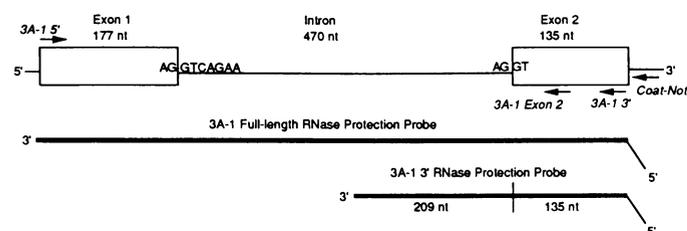
Table 1. In the center column are the oligonucleotide sequences complementary to the gene sequences shown on the right. Upper case letters designate nucleotides complementary to specific gene sequences; lower case letters designate adapter sequences in the PCR oligonucleotide primers used for subcloning. The gene sequences are derived from: pea ribulose 1, 5-bisphosphate carboxylase (rbcS3A, 23); wheat ribulose 1, 5-bisphosphate carboxylase (rbcSws4.3, 24); soybean β-conglycinin β-subunit (β-con, 25); TGMV coat protein (27).

Splicing of the pea *rbcS3A-1* intron in transfected *Nicotiana benthamiana* nuclei

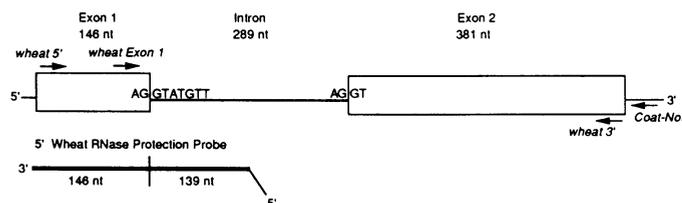
To test the feasibility of the pMON458 transient expression system for studying plant RNA processing, we introduced the first of two introns in the pea (*Pisum sativum* L.) *rbcS3A* gene (23) into pMON458. This intron, designated 3A-1 in this study, is 469 nucleotides in length and contains a pyrimidine-rich sequence upstream of the 3' splice site (23). The 5' splice site is complementary at 7/9 positions to the first nine nucleotides of U1 snRNA that are conserved in mammals, yeasts and plants (28) and known to base-pair with the 5' splice site (4–6).

The *rbcS3A-1* intron and the full length of the two adjacent exons were cloned by PCR amplification into the tomato golden mosaic virus (TGMV) based transient expression vector pMON458 (19) as described in Materials and Methods. In the resulting construct, pMON458-*rbcS3A-1* (Fig. 2A), the *rbcS3A-1* cassette is positioned 8 nucleotides downstream from the TGMV coat protein (AR1) transcription initiation site (Fig. 1C). The 3' end of the *rbcS3A-1* cassette is fused to 31 nucleotides of linker

A. Pea *rbcS3A-1*



B. Wheat *rbcSws4.3*



C. Soybean B-con 132

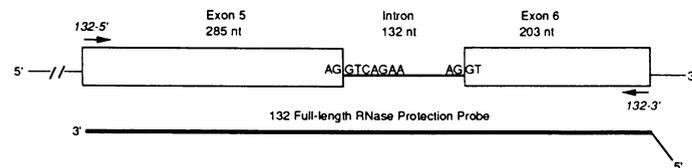


Figure 2. Intron constructs. Shown are the precursor RNAs generated *in vivo* by transcription from the coat protein promoter on pMON458 constructs containing the pea *rbcS3A-1*, wheat *rbcS4.3*, or soybean β -conglycinin 85, 115 and 132 introns. Each transcript is drawn to scale with open boxes designating exon sequences, thick lines designating intron sequences, and thin lines designating linker and TGMV coat protein sequences. Splice site sequences are shown above the exon-intron junctions. The positions and orientation of the oligonucleotide probes outlined in Table 1 are shown with bold arrows above and below the exon sequences. The antisense transcripts used for RNase protection experiments are shown below each transcript with bold lines designating sequences complementary to each of the precursor RNAs and thin lines designating noncomplementary extension sequences present in each antisense transcript used for probe.

DNA, 33 nucleotides of TGMV coat protein coding sequences and the polyadenylation site of the coat protein gene.

Using standard *Agrobacterium* transformation protocols (26), *N. benthamiana* leaf discs were transfected with pMON458-*rbcS3A-1* and leaf discs were collected five days after transfection for DNA and RNA analysis. To verify replication of the transfected vector, total DNA from transfected leaf discs was digested with BglII, electrophoresed on a 1% agarose gel and subjected to Southern blot analysis using a 32 P labeled DNA probe complementary to the TGMV AL1 gene. The resulting Southern blot, shown in Figure 1D, clearly demonstrates that the intact TGMV replicon excised from the pMON458 vector and replicated autonomously in *N. benthamiana* nuclei. We also confirmed that pMON458 constructs containing the *rbcS3A-1* construct inserted in reverse orientation, as well as the second intron and flanking exons of the *rbcS3A* gene (3A-2) and the complete *rbcS3A* gene inserted in both orientations replicated autonomously in plant nuclei. In each case, the two orientations replicated with equal efficiency (Fig. 1D).

To determine if spliced *rbcS3A-1* transcripts accumulated in transfected leaf discs, PCR analysis was initially used to detect transcripts derived from the TGMV replicon. The oligonucleotide primer *Coat-Not*, complementary to the coat protein sequences present at the 3' end of *rbcS3A-1*/coat protein fusion transcript was used to prime cDNA synthesis from total leaf disc RNA. The resulting cDNA served as template for 25 cycles of PCR amplification using the primers designated *3A-1 5'* and *Coat-Not*. This amplification strategy insured that fusion transcripts

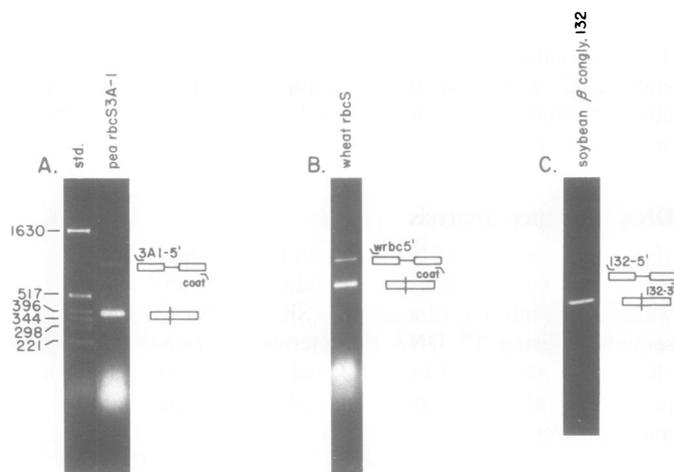


Figure 3. PCR analysis of pre-mRNA splicing in transfected *Nicotiana benthamiana* leaf discs. pMON458 constructs containing the pea *rbcS3A-1*, wheat *rbcS4.3* or soybean β -conglycinin 85, 115 and 132 introns as shown in Fig. 2 were transfected into *N. benthamiana* leaf discs as described in Materials and Methods. Total RNA was isolated five days after transfection and analyzed by PCR amplification using oligonucleotide primers complementary to each of the flanking exons as outlined in Table 1 and Fig. 2. The amplified DNA fragments were electrophoresed on 2% agarose gels and visualized by ethidium bromide staining. The positions of the PCR products corresponding to each of the spliced and unspliced transcripts are designated at the right of each panel. (A) Lane 1, HinfI-cut pBR322 DNA standard; lane 2, 1 μ g RNA isolated from *N. benthamiana* leaf discs transfected with pMON458-*rbcS3A-1* and PCR amplified using the *Coat-Not* and *3A1-5'* oligonucleotide primers. (B) 1 μ g RNA isolated from *N. benthamiana* leaf discs transfected with pMON458-*rbcS4.3* and PCR amplified using the *Coat-Not* and *wheat-5'* oligonucleotide primers. (C) 1 μ g RNA isolated from *N. benthamiana* leaf discs transfected with pMON458-*con85-132* and PCR amplified using the *132-5'* and *132-3'* oligonucleotide primers.

generated by the TGMV expression vector system were detected while endogenous *N. benthamiana* and potential contaminating pea *rbcS* transcripts were not detected. PCR amplification of RNA isolated from leaf discs transfected with the *rbcS3A-1* intron, (Fig. 3A), generated two amplified DNA fragments which correspond to the expected sizes for the unspliced *rbcS3A-1* precursor RNA and the spliced *rbcS3A-1* RNA. Negative control reactions using RNA from leaf discs transfected with the pMON458 vector alone did not produce any detectable PCR amplification products (not shown).

The PCR analysis described above indicated that the *rbcS3A-1* intron was transcribed and spliced in transfected *N. benthamiana* leaf disc nuclei. To define the accuracy of splicing, the PCR product corresponding to the spliced transcript was cloned and sequenced. As shown in Figure 4A, the sequence across the splice junction corresponds to the sequence expected for an accurately spliced *rbcS3A-1* transcript indicating that the *rbcS3A-1* precursor transcript is correctly spliced in transfected leaf disc nuclei.

Splicing of the wheat *rbcS4.3* intron in transfected *N. benthamiana* leaf nuclei

We next examined the splicing of a monocot intron in this transfection system by inserting the wheat (*Triticum aestivum* L.) *rbcS4.3* gene (24) into the pMON458 vector. This gene contains a 289 nucleotide intron analogous to the first intron in the pea *rbcS3A* gene (3A-1). The 3' splice site of the *rbcS4.3* intron is pyrimidine-rich, and the 5' splice site is complementary at 7/9 nucleotides to the 5' end of U1 snRNA.

For our analysis, the wheat *rbcS4.3* intron was cloned with the full-length adjacent exons into pMON458 generating pMON458-*rbcS4.3*. RNA prepared 5 days after transfection was initially analyzed by the PCR amplification strategy described for pea *rbcS3A-1*. As shown in Figure 3B, PCR amplification resulted in the production of two DNA fragments corresponding to the expected products for the *rbcS4.3* precursor and spliced transcripts. The DNA product corresponding to the spliced transcript was cloned and sequenced. The sequence across the

splice junction, (Fig. 4B) indicates that the wheat *rbcS4.3* transcript is accurately spliced in transfected *N. benthamiana* leaf nuclei.

Splicing efficiencies of the pea *rbcS3A-1* and wheat *rbcS4.3* transcripts in transfected *Nicotiana benthamiana* nuclei

RNase protection analysis was used to define the splicing efficiencies of the *rbcS3A-1* and *rbcS4.3* introns. In this analysis, labeled antisense transcripts corresponding to full-length precursor or exon-specific sequences were hybridized with total RNA extracted from transfected leaf discs and subjected to RNase A and T1 digestion. As shown in Figure 5, the abundance and sizes of the protected RNAs define the levels of precursor and spliced RNAs present in each RNA sample. In the pea leaf RNA positive control (Fig. 5A, lane 3), the full-length 3A-1 antisense probe protected 177 nucleotide and 135 nucleotide fragments corresponding to the first and second exons of *rbcS3A* transcripts. Identical fragments were protected in the *rbcS3A-1* transfected *N. benthamiana* leaf disc RNA (Fig. 5A, lane 5). In addition, two higher molecular weight RNAs, representing unspliced pre-mRNA, were detected in the transfected leaf disc RNA. Both RNAs are detected in pea leaf RNA control samples when larger quantities of protected RNA are loaded onto the gel (not shown)

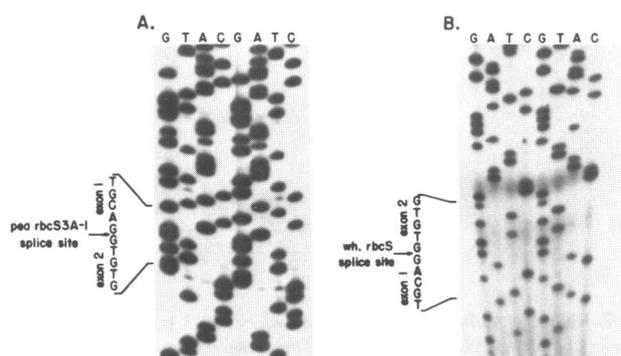


Figure 4. Sequence analysis of spliced transcripts produced in transfected *Nicotiana benthamiana* leaf disc nuclei. (A) The PCR product obtained by amplification of pMON458-*rbcS3A-1* transfected RNA with the *Coat-Not* and *3A1-5'* oligonucleotide primers, which corresponded to the spliced transcript, was cloned and sequenced using the *3A1-exon 2* primer. The sequence across the spliced exon 1/exon 2 junction is indicated at the left. (B) The PCR product obtained by amplification of pMON458-*rbcS4.3* transfected RNA with the *Coat-Not* and *wheat-5'* oligonucleotide primers, which corresponded to the spliced product, was cloned into the Bluescript vector and sequenced using the *wheat-5'* sequencing primer shown in Table 1. The sequence across the spliced exon1/exon2 junction is indicated at the left.

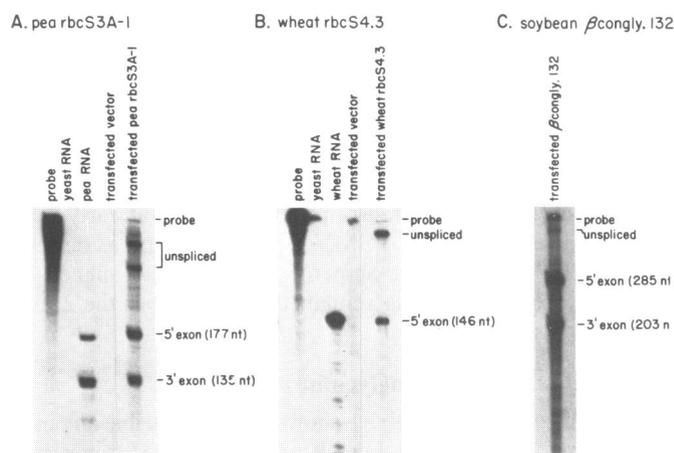


Figure 5. RNase protection analysis of pre-mRNA splicing in transfected *Nicotiana benthamiana* leaf discs. pMON458 constructs containing the pea *rbcS3A-1*, wheat *rbcS4.3* or soybean β -conglycinin 85, 115 and 132 introns were transfected into *N. benthamiana* leaf discs. Total RNA isolated five days after transfection was subjected to RNase protection analysis using ^{32}P -labeled antisense probes (Fig. 2). The protected RNA fragments were electrophoresed on 6% acrylamide, 8.3 M urea denaturing gels and autoradiographed. The positions and sizes of the undigested probe, unspliced transcript, spliced 5' exon and spliced 3' exons are indicated at the right of each panel. (A) RNase protection analysis with the full-length pea *rbcS3A-1* antisense probe. Lane 1, undigested probe; lanes 2–5, RNase protection fragments derived from hybridization of the full-length *rbcS3A-1* antisense probe with: 10 μg total yeast RNA (lane 2), 0.2 μg total pea leaf RNA (lane 3), 10 μg total RNA from *N. benthamiana* leaf discs transfected with pMON458 (lane 4), 1.25 μg total RNA from *N. benthamiana* leaf discs transfected with pMON458-*rbcS3A-1* (lane 5). Two high molecular weight RNAs representing unspliced precursor RNA are designated at the top of lane 5. (B) RNase protection analysis with the 5' exon wheat *rbcS4.3* antisense probe. Lane 1, undigested probe; lanes 2–5, RNase protection fragments derived from hybridization of the 5' exon *rbcS4.3* antisense probe with: 10 μg total yeast RNA (lane 2), 0.2 μg total wheat leaf RNA (lane 3), 10 μg total RNA from *N. benthamiana* leaf discs transfected with pMON458 (lane 4), 1.25 μg total RNA from *N. benthamiana* leaf discs transfected with pMON458-*rbcS4.3* (lane 5). (C) RNase protection fragments derived from hybridization of the full-length soybean β -con 132 antisense probe with 5.0 μg total RNA from *N. benthamiana* leaf discs transfected with pMON458- β con85-132 (lane 1).

indicating that these represent natural RNase cleavage products derived from precursor transcripts probably at AU-rich sequences present in the RNA:RNA hybrids.

When the 3' exon antisense probe for *rbcS3A-1* was used in RNase protection analysis, only the expected spliced second exon and unspliced protected fragments were detected (Fig. 6). Consequently, the *rbcS3A-1* 3' exon probe was used for more quantitative analysis of 3A-1 splicing efficiency. In five independent transfection experiments, we have determined that 80% to 90% (85% average) of the accumulated *rbcS3A-1* RNA was spliced *in vivo* in *N. benthamiana* nuclei.

RNase protection analysis of the wheat *rbcS4.3* spliced products using the *rbcS4.3* 5' exon antisense probe is shown in Figure 5B. In both the wheat leaf RNA positive control and pMON458-*rbcS4.3* transfected leaf disc RNA (lanes 3 and 5), a 146 nucleotide protected fragment corresponding to the correctly spliced first exon of the *rbcS4.3* RNA transcript is readily visible. An additional band corresponding to unspliced transcript is detected in RNA from transfected leaf discs. Five independent transfections and the wheat 5' exon antisense probe indicated that between 69% and 76% (73% average) of the wheat *rbcS4.3* transcripts which accumulated in transfected nuclei were spliced.

Analysis of splicing at various times after transfection

To define the relative splicing efficiencies at various times after transfection, RNase protection experiments were performed on pMON458-*rbcS3A-1* transfected RNA samples isolated between one and nine days after transfection. As shown in Figure 6, precursor and spliced transcripts were first detected two days after transfection and the total transcript levels peaked four days after transfection. The rate of transcript accumulation at early time points reflected the rate of TGMV DNA replication as defined by Southern analysis (not shown). The ratio of spliced to unspliced *rbcS3A-1* splicing remained constant over the 2 to 7 day post-transfection period.

Splicing of soybean β -conglycinin introns

The experiments described above showed that reasonable amounts of unspliced pea *rbcS3A-1* and wheat *rbcS4.3* transcripts accumulate in transfected *N. benthamiana* leaf disc cells. The

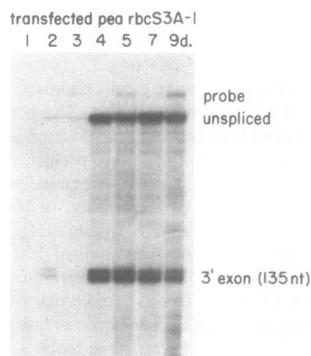


Figure 6. Splicing efficiencies at various times after transfection. Total RNA prepared from *N. benthamiana* leaf discs collected 1, 2, 3, 4, 5, 7, or 9 days after transfection (as designated at the top of the figure) with pMON458-*rbcS3A-1* was analyzed by RNase protection with the 3' exon *rbcS3A-1* antisense probe shown in Fig. 2. The positions of the undigested probe, unspliced transcript and spliced 3' exon are indicated at the right.

low splicing efficiency of the wheat *rbcS4.3* intron (73%) was not unexpected since this monocot intron was inefficiently spliced in transgenic *N. tabacum* plants (15). However, the dicot pea *rbcS3A-1* intron was expected to splice more efficiently than the 85% defined in our initial experiments. To determine whether lower splicing efficiencies are an inherent property of this expression system or of specific introns, leaf discs were transfected with an engineered construct containing the third, fourth and fifth introns of the soybean β -conglycinin gene (25), designated pMON458- β con85-132. PCR analysis with the 132-5' and 132-3' oligonucleotide primers spanning intron 5 (intron 132), indicated that this intron was efficiently spliced in transfected leaf discs (Fig. 3C). PCR analysis with oligonucleotide primers flanking the third and fourth β -conglycinin introns indicated that these introns were spliced with similar efficiencies (not shown). RNase protection analysis of intron 132 splicing using the full-length antisense probe shown in Figure 2C, indicates that intron 132 is accurately spliced from 95% of the accumulated pMON458- β con85-132 transcripts (Fig. 5C). The third and fourth β -conglycinin introns were spliced with comparably high efficiencies (not shown). This series of experiments demonstrated that some introns can be spliced at high efficiency in this

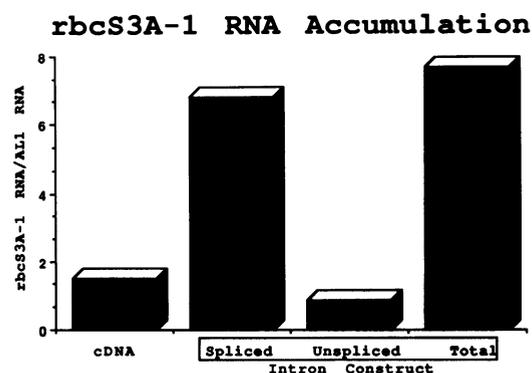
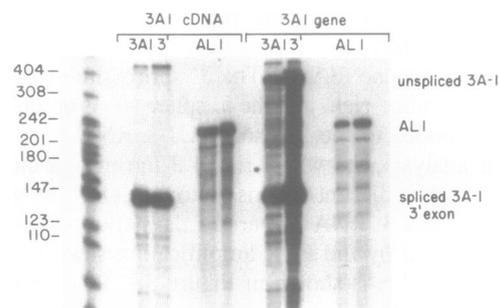


Figure 7. RNA accumulation in *N. benthamiana* leaf discs transfected with the *rbcS3A-1* genomic and cDNA constructs. **(Top)** Total RNA was prepared from *N. benthamiana* leaf discs five days after transfection with the pea *rbcS3A-1* cDNA construct (pMON458-*rbcS3A-1* cDNA) or the *rbcS3A-1* genomic DNA construct (pMON458-*rbcS3A-1*). The RNAs were subjected to RNase protection analysis using the 32 P-labeled 3' exon *rbcS3A-1* antisense (left lanes in each panel) or the AL1 antisense probes (right lanes in each panel) as outlined at the top of the figure. The protected fragments were electrophoresed on a 6% acrylamide, 8.3 M urea denaturing gel and autoradiographed. Each set of paired lanes represents 5 and 10 μ g of input RNA. The positions of the unspliced and spliced 3' exon from the *rbcS3A-1* transcript and the AL1 transcript are indicated at the right. **(Bottom)** After autoradiography, the radioactive bands corresponding to spliced and unspliced transcripts were excised and counted by liquid scintillation spectroscopy. The unspliced and spliced *rbcS3A-1* cpm values were corrected for UMP content and normalized relative to AL1 transcript abundance.

expression system and suggests that the splicing efficiencies defined in this system reflect sequences within particular introns. Additional cotransfection experiments with the *rbcS3A-1*, *rbcS4.3* and β -conglycinin introns demonstrated that each intron maintained its characteristic splicing efficiency in synchronously transfected cells (not shown) indicating that splicing efficiencies defined in this system are highly reproducible under a variety of transfection conditions.

The *rbcS3A-1* intron increases steady-state transcript levels

Several studies have demonstrated that mammalian and plant introns can increase transcript abundance *in vivo* by an undefined mechanism (29–34). To determine if such effects could be detected with the pMON458 transfection system, a cDNA copy corresponding to the *rbcS3A-1* construct without the intron was ligated into pMON458 to produce a pMON458-3A-1cDNA clone. In parallel transfections, the *rbcS3A-1* cDNA and *rbcS3A-1* gene constructs were introduced into *N. benthamiana* leaf discs and transcript accumulation was analyzed by RNase protection (Fig. 7). To control for differences in transfection and replication efficiencies, *rbcS3A-1* transcript levels from gene and cDNA transfections were normalized to the TGMV AL1 transcript levels. Standardization against this transcript indicates that cells containing the gene construct accumulated 4.5-fold more total 3A-1 RNA than those expressing the cDNA and 5.1-fold more spliced 3A-1 RNA than those expressing the intronless cDNA construct.

DISCUSSION

Primary transcripts derived from RNA polymerase II transcription of nuclear genes undergo multiple post-transcriptional RNA processing steps before mature mRNAs are transported to the cytoplasm. The most critical RNA processing events required for expression of translationally competent mRNAs include 5' capping, 3' cleavage and polyadenylation, and, most importantly, intron excision. Although mammalian and yeast RNA splicing pathways have been extensively studied, the nuclear components and sequences required for intron excision in plant nuclei have received far less attention. Dramatic increases in the number of available plant gene sequences have enabled researchers to suggest that plant intron sequences differ significantly from their mammalian and yeast counterparts (16, 35, 36). Most notably, the 5' splice sites in plant introns are loosely conserved, and many introns lack the pyrimidine-rich tract upstream from the 3' splice site (35, 36). In addition, plant introns are very AU-rich (16). The recognition of plant introns appears to occur by a unique mechanism because many plant introns are not properly excised from precursor transcripts in mammalian RNA splicing extracts (13, 14) and mammalian introns are not accurately spliced in plant nuclei (14, 37). These differences between eukaryotic RNA processing systems exist even within the plant kingdom since some monocot introns are inefficiently spliced from precursor transcripts expressed in transgenic dicot nuclei (15). Although one might expect the basic intron cleavage and ligation reactions to be identical in mammalian, yeast and plant nuclei, it is apparent that each of these organisms has evolved special mechanism(s) for intron recognition.

The populations of splicesomal snRNAs, which are integrally involved in intron recognition, are substantially more complex in plant nuclei than in mammalian nuclei and vary between the monocot wheat and the dicot pea (38, 39) suggesting that the

monocot and dicot RNA splicing machineries have diverged. The novel properties of plant introns and the complexity of the snRNA components involved in their excision make thorough studies of the intron sequence requirements essential for understanding the mechanism of plant intron splicing. Until now, such studies have been limited by the availability of fast, efficient plant expression systems that produce abundant pre-mRNA transcripts *in vivo*. In one available system, the integration of test constructs into transgenic plants requires extended periods of time which dramatically limit the number of constructs that can be analyzed. In the other system, which involves transient expression of precursor RNAs from nonreplicating vectors in protoplasts, analysis is much faster but less efficient and experimental interpretations may be complicated by the physiological state of the protoplasts. Thus, new plant transient expression systems are needed for analysis of intron splicing requirements.

In this paper, we have investigated the usefulness of a TGMV geminivirus-based transient expression system (19) for *in vivo* analysis of plant intron splicing. Using this system, intron constructs have been expressed in *N. benthamiana* leaf nuclei from multi-copy autonomous TGMV replicons. PCR analysis, DNA sequence analysis and S1 nuclease analysis (not shown) have demonstrated that the pea *rbcS3A-1*, wheat *rbcS4.3* and soybean β -conglycinin 132 introns are accurately excised from pre-mRNAs expressed in transfected leaf nuclei (Figs. 3 and 4). Quantitative RNase protection analysis demonstrated that the β -conglycinin 132 intron was accurately spliced from greater than 95% of the accumulated transcripts while the pea *rbcS3A-1* and wheat *rbcS4.3* introns were spliced from 85% and 73% of accumulated transcripts, respectively (Fig. 5). Analysis of spliced and unspliced transcripts at various times after transfection demonstrated that pea *rbcS3A-1* transcripts can be readily detected by RNase protection analysis (Fig. 6) or PCR analysis (not shown) two days after transfection. After this, between two and nine days post-transfection, transcripts accumulated at higher levels proportional to the number of TGMV replicons present in each cell. In spite of the variations in transcript abundance, the proportion of spliced transcript remained constant between two and seven days after transfection. Thus, it is clearly evident that the high expression levels of precursor RNA in this system do not exceed the capacity of the intron splicing machinery present in these nuclei. We have concluded that the replicating viral-based vector utilized in this paper provides the basis for a highly efficient and accurate expression system for studying plant RNA processing.

The splicing efficiencies that we have defined for the ribulose biphosphate carboxylase introns (73–85%) and β -conglycinin introns 3, 4, and 5 (>95%) in the geminivirus transient expression system are higher than the splicing efficiencies generally defined for plant introns in transient protoplast systems. In studies involving the expression of monocot introns in CAT or GUS constructs in electroporated maize protoplasts, 10–50% of the expressed pre-mRNA that accumulated remained unspliced (18). PEG-transfected *N. plumbaginifolia* protoplasts, three introns from the french bean phaseolin gene were spliced at 57–67% efficiency while an intron from the soybean leghemoglobin gene spliced at 78% efficiency (16). The 67% splicing efficiency reported for phaseolin intron 4 in protoplasts dramatically contrasts with the high (>95%) splicing efficiency defined in our transfected leaf nuclei for the equivalent intron from the highly homologous β -conglycinin gene. The proportionately higher splicing efficiencies that we have obtained

in the geminivirus-transfected nuclei presumably reflect the physiological integrity of the splicing machinery in the leaf disc nuclei. It may also reflect the fact that, in the geminivirus expression system, precursor transcripts are generated from autonomously replicating vectors assembled in chromatin complexes (40). Transcription from such complexes may favorably affect the secondary structures assumed by each precursor transcript. Regardless of the basis for increased splicing efficiencies, the optimal efficiencies obtained in this transient expression system allow for an accurate definition of critical intron recognition sequences.

Both the pea *rbcS3A-1* and the wheat *rbcS4.3* introns are accurately excised from precursor transcripts in transfected *N. benthamiana* leaf nuclei (Fig. 4). As reported by Keith and Chua (15) using intron constructs similar to those described here, the pea *rbcS3A-1* intron was spliced from approximately 100% of the transcripts which accumulated in transgenic *N. tabacum* plants and the wheat *rbcS4.3* intron was spliced from approximately 70% of the accumulated transcripts. From this analysis and analysis of maize *Adh-1* intron 6, Keith and Chua suggested that dicot nuclei inefficiently process some monocot pre-mRNAs. In agreement with this analysis, our studies with the viral expression vectors have demonstrated that the wheat *rbcS4.3* intron was spliced from 73% of the transcripts which accumulated in the *N. benthamiana* leaf nuclei (Fig. 5). In contrast, only 85% of the pea *rbcS3A-1* transcripts expressed from autonomously replicating plasmids in transfected *N. benthamiana* nuclei were spliced (Fig. 5). The lower splicing efficiency of the pea *rbcS3A-1* intron is not characteristic of this expression system since each of the introns in a transcript containing three soybean β -conglycinin introns can be spliced at greater than 95% efficiency (Fig. 5). In addition, as demonstrated in Fig. 6, large variations in *rbcS3A-1* transcript abundance occur without significantly changing the observed splicing efficiency. Three possible explanations exist for the lower apparent splicing efficiency of the pea *rbcS3A-1* intron in *N. benthamiana* leaf discs compared with that observed in transgenic *N. tabacum* plants. One potential explanation is that the *rbcS3A-1* transcript levels saturate the capacity of the splicing machinery in the transfected nuclei. This appears unlikely because three soybean β -conglycinin introns were spliced at optimal efficiency from a transcript that is as abundant as the pea *rbcS3A-1* transcript relative to the *AL1* control transcript (not shown). One caveat is that splicing of the *rbcS3A-1* intron may require additional recognition factors not utilized in the splicing of the β -conglycinin introns which may be limiting in *N. benthamiana* leaf nuclei. In a second explanation, the low splicing efficiency apparent in the transient expression system may actually reflect the general instability of unspliced *rbcS3A-1* precursor RNAs in transgenic versus transient expression systems. The third and, potentially, most interesting explanation ensues from the fact that the transfected leaf discs used in this geminivirus expression system are cultured on media containing 3% sucrose. Sucrose, which negatively regulates photosynthetic gene expression at a transcriptional level (41), conceivably causes photosynthetic cells to dedifferentiate into non-photosynthetic cells which may contain limiting leaf-specific splicing factors.

In an increasing number of cases, expression of mammalian and plant genes is enhanced 2- to 500-fold by the presence of an intron within the transcription unit (18, 29–31, 42). For some mammalian introns, this effect is mediated by transcriptional enhancers located within the intron (29), while in others the

enhancement is mediated by elements directly affecting the level of mRNA (33). With one exception (32), all of the reported instances of plant intron enhancement occur when monocot introns are introduced into monocot cells (18, 31, 34, 42). Despite expression increases between 5- and 100-fold, the mechanism(s) of plant intron enhancement is unknown. Because the TGMV vector provides a potential expression system for experimentally defining this mechanism in dicot cells, we analyzed the enhancement properties of the pea *rbcS3A-1* intron *in vivo* in *N. benthamiana* leaf discs. Expression of a construct containing the pea *rbcS3A-1* intron resulted in accumulation of 5.1-fold more RNA than expression of an intronless *rbcS3A-1* cDNA construct (Fig. 7). This level of enhancement is comparable to the 5-fold enhancement reported for expression of the petunia *rbcS* first intron in transgenic tobacco plants (32) and contradicts previous reports suggesting that the dicot introns do not enhance expression in transgenic *N. tabacum* plants (34, 43).

In summary, we have defined the utility of a TGMV-based autonomously replicating vector for analysis of plant RNA processing in *N. benthamiana* leaf nuclei. Our studies have demonstrated that monocot and dicot introns can be accurately excised in this system and that some introns are spliced at essentially optimal efficiency. Other introns are spliced at slightly reduced efficiencies reflecting particular intron sequences rather than an ineffective splicing machinery. In this system, the differences in the splicing efficiencies of the monocot and dicot *rbcS* introns do not appear to be as extensive as previously documented in transgenic plants (15). High level expression of precursor transcripts in the transient expression system clearly indicates that tobacco leaf disc nuclei are capable of effectively and accurately processing monocot introns with efficiencies similar to dicot introns. Quantitative splicing analysis of other introns will determine whether leaf nuclei discriminate at all in the processing of monocot and dicot introns or whether the respective precursor transcripts are differentially stabilized in transgenic nuclei. The latter case would make it appear as though the monocot introns are spliced at reduced efficiencies in transgenic plants. This and similar systems provide excellent alternatives to protoplast transient expression. Such vectors, in combination with transgenic plant studies, should allow us to answer many intriguing plant RNA processing questions.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge Dr. Stephen Rogers and Monsanto for providing the pMON458 vector. We thank Dr. Nam-Hai Chua for providing the pea *rbcS3A* and wheat *rbcS4.3* clones and Carl Sung for generating the pMON458-3A-1 construct. We are indebted to Brian A. Hanley for establishing the TGMV system in our laboratory and for his extreme motivation and inspiration. This work was supported by grant GM39025 RO1 from the National Institutes of Health.

REFERENCES

1. Padgett, R.A., Grabowski, P.J., Konarska, M.M., Seiler, S. and Sharp, P.A. (1986) *Ann. Rev. Biochem.*, **55**, 1119–1150.
2. Keller, E.B. and Noon, W.A. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 7417–7420.
3. Steitz, J.A., Black, D.L., Gerke, V., Parker, K.A., Kramer, A., Frensdewey, D., and Keller, W. (1988) In Birnstiel, M.L. (ed.) *Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles*. Springer-Verlag, Berlin, pp. 115–154.

4. Mount, S.M., Petterson, I., Hinterberger, M., Karmas, A. and Steitz, J.A. (1983) *Cell*, **33**, 509–518.
5. Black, D.L., Chabot, B. and Steitz, J.A. (1985) *Cell*, **42**, 737–750.
6. Zhuang, Y. and Weiner, A.M. (1986) *Cell*, **46**, 827–835.
7. Zhuang, Y. and Weiner, A.M. (1989) *Genes & Dev.*, **3**, 1545–1552.
8. Chabot, B., Black, D.L., LeMaster, D.M. and Steitz, J.A. (1985) *Science*, **230**, 1344–1349.
9. Black, D.L. and Steitz, J.A. (1986) *Cell*, **46**, 697–704.
10. Berget, S.M. and Robberson, B.L. (1986) *Cell*, **46**, 691–696.
11. Brown, J.W.S., Feix, G. and Frendeway, D. (1986) *EMBO J.*, **5**, 2749–2758.
12. Hartmuth, K. and Barta, A. (1986) *Nucl. Acids Res.*, **14**, 7513–7527.
13. Van Santen, V. and Spritz, R. (1987) *Gene*, **56**, 253–265.
14. Wiebauer, K., Herrero, J.-J., and Filipowicz, W. (1988) *Mol. Cell Biol.*, **8**, 2042–2051.
15. Keith, B. and Chua, N.-H. (1986) *EMBO J.*, **5**, 2419–2425.
16. Goodall, G.J. and Filipowicz, W. (1989) *Cell*, **58**, 473–483.
17. Goodall, G.J., Wiebauer, K. and Filipowicz, W. (1990) *Methods in Enzy.*, **181**, 148–161.
18. Luehrsens, K.R. and Walbot, V. (1990) *Mol. Gen. Genet.*, In press.
19. Hanley-Bowdoin, L., Elmer, J.S. and Rogers, S.G. (1988) *Nucl. Acids Res.*, **16**, 10511–10529.
20. Rogers, S.G., Bisaro, D.M., Horsch, R.B., Fraley, R.T., Hoffmann, N.L., Brand, L., Elmer, J.S. and Lloyd, A.M. (1986b) *Cell*, **45**, 593–600.
21. Elmer, J.S., Brand, L., Sunter, G., Gardiner, W.E., Bisaro, D.M. and Rogers, S.G. (1988) *Nucleic Acids Res.*, **16**, 7043–7060.
22. Gardiner, W.E., Sunter, G., Brand, L., Elmer, J.S., Rogers, S.G. and Bisaro, D.M. (1988) *EMBO J.*, **7**, 899–904.
23. Fluhr, R., Moses, P., Morelli, G., Coruzzi, G. and Chua, N.-H. (1986) *EMBO J.*, **5**, 2063–2071.
24. Broglie, R., Coruzzi, G., Lamppa, G., Keith, B. and Chua, N.-H. (1983) *Biotech.*, **1**, 55–61.
25. Schuler, M.A., Schmitt, E.S. and Beachy, R.N. (1982) *Nucl. Acids Res.*, **10**, 8225–8244.
26. Rogers, S.G., Horsch, R.B. and Fraley, R.T. (1986a) *Methods in Enzymology*, **118**, 627–640.
27. Hamilton, W.D.O., Stein, V.E., Coutts, R.H.A. and Buck, K.W. (1984) *EMBO J.*, **3**, 2197–2205.
28. Reddy, R. and Busch, H. (1988) In Birnstiel, M.L. (ed.) *Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles*. Springer-Verlag, Berlin, pp 1–37.
29. Brinster, R.L., Allen, J.M., Behringer, R.R., Gelinas, R.E. and Palmiter, R.D. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 836–840.
30. Buchman, A.R. and Berg, P. (1988) *Mol. Cell Biol.*, **8**, 4395–4405.
31. Callis, J., Fromm, M. and Walbot, V. (1987) *Genes and Devel.*, **1**, 1183–1200.
32. Dean, C., Favreau, M., Bond-Nutter, D., Bedbrook, J. and Dunsmuir, P. (1989) *The Plant Cell*, **1**, 201–208.
33. Huang, M.T.F. and Gorman, C.M. (1990) *Nucl. Acids Res.*, **18**, 937–947.
34. Tanaka, A., Satoru, M., Shozo, O., Kyozyuka, J., Shimamoto, K. and Nakamura, K. (1990) *Nucl. Acids Res.*, **18**, 6767–6770.
35. Brown, J.W.S. (1986) *Nucl. Acids Res.*, **14**, 9549–9559.
36. Hanley, B.A. and Schuler, M.A. (1988) *Nucl. Acids Res.*, **16**, 7159–7176.
37. Barta, A., Sommergruber, K., Thompson, D., Hartmuth, K., Matzke, M.A. and Matzke, A.J.M. (1986) *Plant Mol. Biol.*, **6**, 347–357.
38. Kiss, T., Toth, M. and Solymosy, F. (1985) *Eur. J. Biochem.*, **152**, 259–266.
39. Egeland, D.B., Sturtevant, A.P. and Schuler, M.A. (1989) *The Plant Cell*, **1**, 633–643.
40. Abouzid, A.M., Frischmuth, T. and Jeske, H. (1988) *Mol. Gen. Genet.*, **212**, 252–258.
41. Sheen, J. (1990) *The Plant Cell*, **2**, 1027–1038.
42. McElroy, D., Zhang, W., Cao, J. and Wu, R. (1990) *The Plant Cell*, **2**, 163–171.
43. Kuhlemeier, C., Fluhr, R. and Chua, N.-H. (1988) *Mol. Gen. Genet.*, **212**, 405–411.
44. Sunter, G., Gardiner, W.E. and Bisaro, D.M. (1989) *Virology*, **170**, 243–250.