PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Plant pre-mRNA Splicing and Splicing Components

John W. S. Brown, Craig G. Simpson, Gordon G. Simpson, Andrew D. Turnbull-Ross and Gillian P. Clark

Phil. Trans. R. Soc. Lond. B 1993 **342**, 217-224
doi: 10.1098/rstb.1993.0150

Plant pre-mRNA splicing and splicing components

JOHN W. S. BROWN, CRAIG G. SIMPSON, GORDON G. SIMPSON, ANDREW D. TURNBULL-ROSS AND GILLIAN P. CLARK

Cell and Molecular Genetics Department, Scottish Crop Research Institute, Invergowrie, Dundee DD2 5DA, U.K.

SUMMARY

Pre-mRNA splicing or the removal of introns from precursor messenger RNAs depends on the accurate recognition of intron sequences by the plant splicing machinery. The major components of this machinery are small nuclear ribonucleoprotein protein particles (snRNPs) which consist of snRNAs and snRNP proteins. We have analysed various aspects of intron sequence and structure in relation to splice site selection and splicing efficiency and we have cloned snRNA genes and a gene encoding the snRNP protein, U2B″. In the absence of an *in vitro* splicing system for plants, transient expression in protoplasts and stable plant transformations have been used to analyse splicing of intron constructs. We aim to address the function of the UsnRNP-specific protein, U2B″, via the production of transgenic plants expressing antisense U2B″ transcripts and epitope-tagged U2B″ protein. In addition, we have cloned genes encoding other proteins which potentially interact with RNA, such as RNA helicases, and strategies involving transgenic plants are being developed to analyse their function.

1. INTRODUCTION

Most eukaryotic protein-coding genes contain intervening sequences (introns) which must be removed from precursor messenger RNA (pre-mRNA) transcripts before the resultant mRNAs can be translated. Pre-mRNA splicing is one of a number of processing events which includes capping, 3′ end cleavage and polyadenylation, transport and ultimately translation. An understanding of the mechanisms involved in pre-mRNA splicing in plants is essential, not only because it is a fundamental process in plant gene expression but also because it represents one level at which gene expression can be regulated (Smith *et al.* 1989). Similarly, an appreciation of the intermolecular interactions involved in this process is required to study the splicing process and will provide information which will aid the analysis of other post-transcriptional processing events such as precursor ribosomal RNA processing.

Accuracy of pre-mRNA splicing is clearly essential to the production of correctly spliced mRNAs, maintaining the correct open reading frame and polypeptide production. Accurate intron removal depends on mechanisms for the definition of intron sequences within a pre-mRNA and for recognition of the 5′ and 3′ splice sites. These are mediated by conserved sequences in the intron, and by spliceosomal U-type small nuclear ribonucleoprotein particles (UsnRNPs) and protein factors which interact to form a large complex called the spliceosome (Green 1991; Guthrie 1991; Ruby & Abelson 1991). The study of splicing and splicing components is far more advanced in vertebrate and yeast systems than in plants due to the availability of *in vitro* splicing systems. The overall similarities in key sequences and components between all three systems suggests that much of the current information on animal and yeast splicing is relevant to plant intron splicing. Nevertheless, important differences exist in the plant system making its analysis essential to understanding gene expression in plants. Furthermore, the ease of production of transgenic plants, the efficacy of antisense gene regulation in plants and the evolving technology in terms of inducible promoters may provide unique opportunities for analysing aspects of splicing *in vivo* in plants which will provide more generally applicable information.

2. PRE-mRNA SPLICING IN PLANTS

Besides the importance of splicing as a basic eukaryotic cellular process, splicing differences, intron-enhanced expression and alternative splicing phenomena all need to be addressed in plant systems. Firstly, there are differences in splicing of animal and plant introns as exemplified by the lack of, or inaccurate and poor splicing of animal introns in plant cells (Barta *et al.* 1986; Van Santen & Spritz 1987; Wiebauer *et al.* 1988) and plant introns in the animal *in vitro* splicing system (Wiebauer *et al.* 1988). More surprisingly there is a difference in splicing between monocotyledonous (monocot) and dicotyledonous (dicot) plants (Keith & Chua 1986; Goodall & Filipowicz 1991). The poor efficiency of splicing of monocot introns in dicot cells has implications for expression of intron-containing monocot gene constructs in transgenic dicot plants. Secondly, the phenomenon of intron-enhanced expression is well documented for a number of monocot introns (Callis *et al.* 1987) and inclusion of such

Phil. Trans. R. Soc. Lond. B (1993) **342**, 217–224
Printed in Great Britain

introns into transformation vectors is being used to ensure high levels of gene expression (Sinibaldi & Mettler 1992, and therein). The mechanism by which enhancement of expression occurs is not understood but again clearly has implications for transgenic plant production. Finally, although at present little is known about alternative splicing in plants, this process has perhaps the greatest potential to influence our current way of thinking about regulation of plant gene expression and function. In alternatively spliced systems more than one mRNA is produced from a single pre-mRNA by selection of different combinations of splice sites, and the resultant mRNAs are translated into functionally different but related proteins. Alternative splicing pathways are well documented in animal systems where most are involved in regulating development and differentiation through developmental-stage or tissue-specific regulation of alternative splicing pathways (reviewed by Smith *et al.* 1989). In plants, there are, as yet, only a few examples of alternative splicing (for example, in spinach and *Arabidopsis* rubisco activase; Werneke *et al.* 1989) and little information is available on the functional significance of the alternative products. If alternative splicing is also an important factor in the control of plant development and differentiation, an understanding of intron sequence and structure, the factors involved in splice site recognition and spliceosome assembly will be a pre-requisite to studying the regulation of differentially spliced pathways.

3. INTRON SPLICING SIGNALS

Plant introns are fundamentally similar to yeast and animal introns although clear differences exist which imply modified or novel intron recognition systems. Plant splice site consensus sequences resemble those of animals being extremely variable around the prerequisite :GU and AG: dinucleotides at the 5′ and 3′ splice sites respectively (Brown 1986; Hanley & Schuler 1988; Goodall *et al.* 1991; Sinibaldi & Mettler 1992). In addition, plant and animal introns do not contain the absolutely conserved branchpoint sequence, UACUAAC, found in yeast although animal introns contain a related sequence with the consensus CURAY. Branchpoints have not yet been functionally defined in plant introns although putative branchpoint sequences with the same consensus as in animals can often be identified (Brown 1986). Plant introns also have a generally less pronounced polypyrimidine tract upstream of the 3′ splice site which is an important splicing signal in animals. Subtle variations between monocot and dicot splice sites have been proposed (Hanley & Schuler 1988) but their relevance to splicing remains to be demonstrated in the light of the variability observed in plant intron splice sites, in general, and in splice sites of different introns within the same pre-mRNA transcript, in particular. Finally, plant introns are generally AU-rich and this property, and the associated difference between AU content of intron and exon sequences, are potentially very important factors in intron definition in plants (Wiebauer *et al.* 1988; Goodall & Filipowicz 1989, 1991). The above differences in splicing signals and AU richness probably explain the difference in splicing between vertebrate and plant introns and the poor efficiency of splicing of monocot introns in dicot cells (Goodall & Filipowicz 1991).

4. COMPONENTS OF THE SPLICEOSOME

The major components of the spliceosome are four UsnRNPs, U1, U2, U4/U6 and U5 which consist either of a single U-type small nuclear RNA (UsnRNA) or two UsnRNAs in the case of U4/U6, complexed with a number of proteins. Each of the four UsnRNAs is associated with at least eight common or core proteins and varying numbers of UsnRNP-specific proteins (reviewed by Lührmann *et al.* 1990). The spliceosome assembles in a stepwise fashion around the pre-mRNA which during transcription has been complexed with hnRNP proteins, some of which at least, associate in a sequence-specific manner (Bennett *et al.* 1992). The association of U1snRNP with the 5′ splice site and U2snRNP with the branchpoint sequence leads to formation of a pre-splicing complex and involves the actions of a number of splicing factors (reviewed by Rosbash & Seraphin 1991). The subsequent association of the U4/U6 and U5snRNPs forms the active spliceosome in which exon sequences are cleaved and ligated and the intron released. Throughout this dynamic assembly process intermolecular interactions are occurring at different levels. In the first place, a number of RNA–RNA base pairing interactions are involved between different UsnRNAs and between UsnRNAs and intron and flanking exon sequences in the pre-mRNA (Wassarman & Steitz 1992; Madhani & Guthrie 1992, and references therein). The dynamic formation and dissolution of these RNA interactions may be controlled by a number of RNA helicase splicing factors which have been described mainly in yeast (reviewed by Guthrie 1991; Ruby & Abelson 1991). In addition, to this type of RNA–protein interaction, many splicing components interact with or bind to the pre-mRNA (hnRNP proteins, splicing factors) or the various UsnRNAs (core or UsnRNP-specific proteins). Some hnRNP and UsnRNP proteins now provide model systems for studying RNA–protein binding (reviewed by Kenan *et al.* 1991). Finally, besides RNA–protein interactions, protein–protein interactions are clearly important in UsnRNP assembly and overall spliceosome assembly. To understand the sheer complexity of the spliceosome in terms of number of RNA and protein components involved and how they associate with one another spatially and temporally to build the spliceosomal catalytic complex, is a major challenge in biochemistry and molecular biology. Again the advantages of *in vitro* systems, the abundance of UsnRNPs in animal cells and the power of yeast as a genetic tool has allowed immense progress to be made in these systems.

In plants, analysis of UsnRNA genes has demonstrated far greater sequence variation than their animal counterparts (Vankan & Filipowicz 1988; Waugh *et al.* 1991*a*,*b*; Hanley & Schuler 1991; Abel

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

et al. 1989; Vaux *et al.* 1992; Leader *et al.* 1993). Nevertheless, in terms of the RNA–RNA interactions involved in splicing, either no variation in these important regions is observed or variation is complementary and maintains the base-pairing interaction. The overall similarity of splice site consensus sequences and UsnRNA sequences, and the conservation of specific UsnRNA regions known to be involved in splicing in either vertebrate or yeast systems suggests that many aspects of spliceosomal assembly and the splicing reaction are conserved in plant splicing. While the UsnRNA components have been well characterized in plants very little is known about the protein components of the spliceosome. Biochemical characterization of purified UsnRNPs from *Vicia faba* has identified putative plant UsnRNP protein analogues of human proteins on the basis of protein mobility and immunological identity (Palfi *et al.* 1989). More recently immunological cross reactivity of a plant protein has been demonstrated using antibodies against a yeast U5snRNP protein (Kulesza *et al.* 1993). The only plant UsnRNP core protein genes isolated to date are an alfalfa cDNA with some homology to the human E protein gene (Hirt *et al.* 1992) and the rice D and E protein partial cDNAs isolated as part of the rice cDNA sequencing project. Finally, the only UsnRNP-specific protein genes to have been isolated from plants so far are those encoding the U2snRNP-specific protein, U2B″ from potato (Simpson *et al.* 1991), a partial U2B″ cDNA from rice (Uchimiya *et al.* 1992) and the partial U1–70 kDa protein gene from *Arabidopsis* (Reddy *et al.* 1992). Potato U2B″ remains the best characterized plant UsnRNP protein and as an RNA-binding protein is providing an important evolutionary comparison to model systems in animals.

5. SPLICING ANALYSIS OF PLANT INTRONS

(*a*) *Expression vector for pre-mRNA splicing analysis*

In the absence of a plant *in vitro* splicing system we and others have developed methods of analysing splicing either in transient protoplast systems (Goodall *et al.* 1991; Sinibaldi and Mettler 1992; Waigmann and Barta 1992; Simpson and Brown 1993) or in plant cells following *Agrobacterium*-mediated transfer of an intron-containing gemini virus based vector (McCullough *et al.* 1991). We have developed a vector for intron expression in protoplast systems from which inserts are readily transferred to transformation vectors for production of transgenic plants, and have applied a sensitive reverse transcription-polymerase chain reaction (RT–PCR) method to the analysis of intron splicing. While transient protoplast methods are more amenable to the analysis of large numbers of intron constructs, transgenic plants can be used to confirm findings with the transient system. The intron-expression vector consists of the intronless coding region of a maize zein storage protein gene from pMS2 (Langridge *et al.* 1985) cloned into a derivative of the plant expression vector, pDH51 (Pietrzak *et al.*

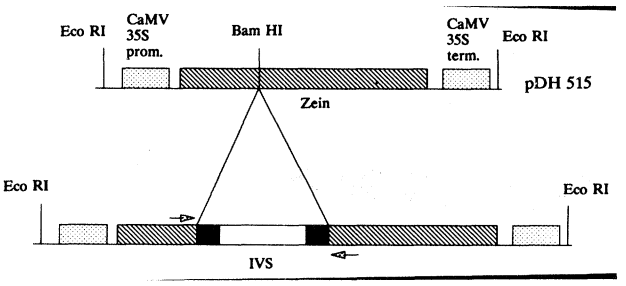


Figure 1. Intron-expression vector. Plant introns and adjacent exons are introduced into a unique *Bam*HI site in the zein gene and transcribed from the CaMV 35S promoter. Bordering *Eco*RI sites are utilized to transfer the cassette to plant transformation vectors.

1986) between the Cauliflower Mosaic Virus (CaMV) 35*S* promoter and termination/polyadenylation sequence to generate plasmid pDH515. This plasmid contains a unique *Bam*HI site in the zein gene at position +170 relative to the translation initiation codon into which intron sequences can be readily introduced and transcribed as part of a full length protein coding pre-mRNA transcript (figure 1). Intron-containing plasmids (30 μg) are routinely

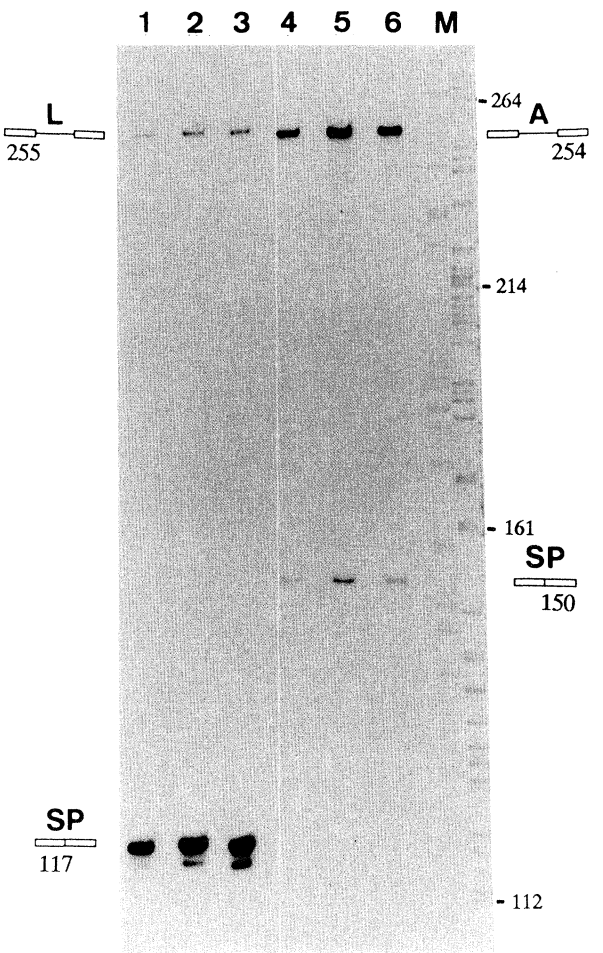


Figure 2. RT–PCR analysis of splicing in tobacco protoplasts. Unspliced (L and A) and spliced products (SP) of three separate transfections of the legumin intron (lanes 1–3) and amylase intron (lanes 4–6) are aligned with a DNA sequencing marker.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

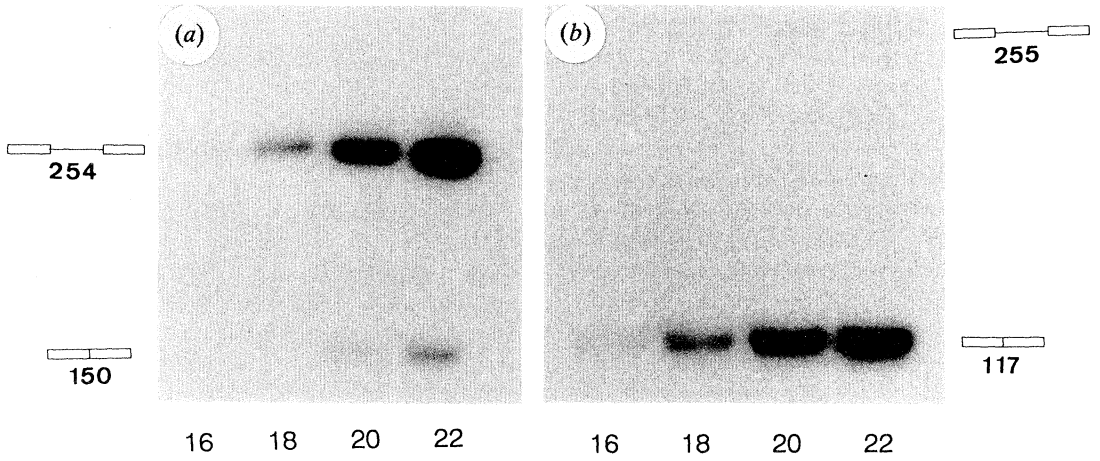


Figure 3. Autoradiography of RT–PCR products for quantification. Unspliced and spliced products of the amylase construct (*a*) and legumin construct (*b*) from aliquots of the RT–PCR reaction after 16 to 22 cycles were detected by autoradiography following separation on a polyacrylamide gel.

transfected into protoplasts (2×10^5) prepared from 4–6 week old leaf material of *Nicotiana tabacum* var. Xanthii using polyethylene glycol (Guerineau *et al.* 1991). Alternatively, the whole intron-containing cassette can be isolated as an *Eco*RI fragment (figure 1) for introduction into pBin19 (Bevan 1984) or other appropriate transformation vectors for the production of transgenic potatoes.

Zein genes are expressed exclusively in the developing endosperm of maize and primers to sequences within the zein coding region bordering the site of intron insertion were selected for RT–PCR and routinely gave no background amplification signals with control untransfected tobacco leaf protoplast RNA or potato leaf RNA. The RT–PCR method of Golde *et al.* (1990), in which one of the oligonucleotide primers is end-labelled with ^{32}P producing labelled RT–PCR products corresponding to unspliced and spliced RNAs, provided an accurate, sensitive and quantifiable method of RNA transcript analysis (Simpson & Brown 1992) for examining splicing of intron constructs.

(*b*) *Splicing of plant introns in tobacco protoplasts and transgenic potato*

Constructs to examine intron recognition, splice site selection and splicing efficiency between monocots and dicots were based on a pea legumin J intron (dicot) and a wheat amylase 33 intron (monocot) and have been described previously (Brown *et al.* 1986). RT–PCR analysis of three separate transfections of expression plasmids containing the monocot wheat amylase intron (pA) and the dicot pea legumin intron (pL) in tobacco protoplasts are shown in figure 2. Both introns were accurately spliced as evidenced by the size of the spliced products (figure 2) and sequencing of cloned spliced products. While both introns have similar 5′ splice sites (amylase–ATC:GTAAGT, legumin–AGA:GTAAGT) and 3′ splice sites (amylase–GCAG:T, legumin–GCAG:G), matching plant consensus sequences, their AU contents are substantially different (amylase, 55% AU; legumin, 75% AU). Although these naturally occurring introns

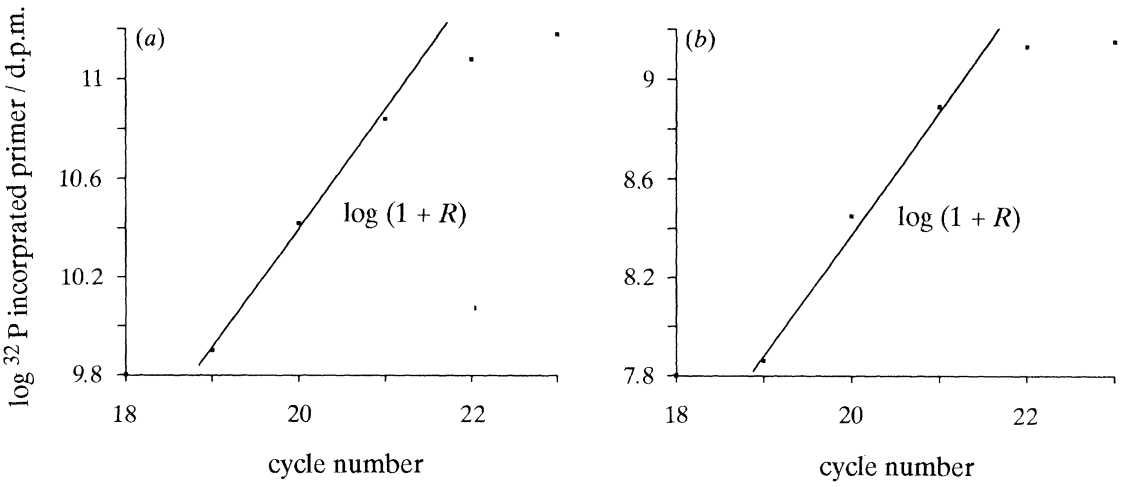


Figure 4. Plot of log counts against PCR cycle number for amylase unspliced (*a*) and spliced (*b*) transcripts. The efficiency of *Taq* polymerase is calculated by the slope of the line in the linear region of the graph and used to calculate the amount of cDNA prior to PCR as described by Golde *et al.* (1990).

contain other: GU and AG: dinucleotides, no evidence for the activation of cryptic splice sites was obtained.

To obtain an estimate of splicing efficiency for each intron construct, it was necessary to quantify PCR products. Quantification of the amount of each RT–PCR product was determined as described by Golde *et al.* (1990) by measuring the incorporation of end-labelled 5′ primer into the PCR product(s) over a range of PCR cycles. Aliquots of the PCR reaction were removed after specific cycles (typically between 16 to 25 cycles), separated on polyacrylamide gels, visualized by autoradiography (figure 3) and bands excised for Cerenkhov counting. Plotting of log counts against cycle number typically showed a linear relationship over cycle number ranges of 16 to 21 (figure 4). From the slope of the line in this region it is possible to calculate the efficiency of the *Taq* polymerase for amplification of each transcript and thereby the amount of cDNA before PCR (Golde *et al.* 1990). Splicing efficiency was determined by calculating the amount of spliced product as a percentage of the total from three separate transfections (table 1). Although both introns were accurately spliced, the legumin intron was spliced with an efficiency of $82 \pm 7\%$ while the monocot amylase intron was poorly spliced with $2 \pm 1\%$ efficiency.

The same intron-expressing cassettes can be introduced into transgenic plants and analysed by the RT–PCR method. A number of intron constructs have now been analysed in both plant protoplasts and transgenic plants. For the most part, results of splicing of the

Table 1. *Splicing efficiency*

(*Taq* efficiency and the amount of cDNA at time zero ($cDNA_0$ were calculated using the formula of Golde *et al.* (1990).)

		transfection 1			transfection 2			transfection 3			
plasmid	transcript	*Taq* efficiency	$cDNA_0$ (d.p.m. $\times 10^{-3}$)	splicing efficiency %	*Taq* efficiency	$cDNA_0$ (d.p.m. $\times 10^{-3}$)	splicing efficiency %	*Taq* efficiency	$cDNA_0$ (d.p.m. $\times 10^{-3}$)	splicing efficiency %	Mean splicing efficiency %
pL	unspliced	0.68	104	68	0.8	46	93	0.5	2330	86	82 ± 7
	spliced	0.86	223		0.81	628		0.57	14570		
pA	unspliced	0.52	4982	0.7	0.6	2817	4.6	0.55	3044	0.9	2 ± 1
	spliced	0.74	37		0.68	135		0.75	29		

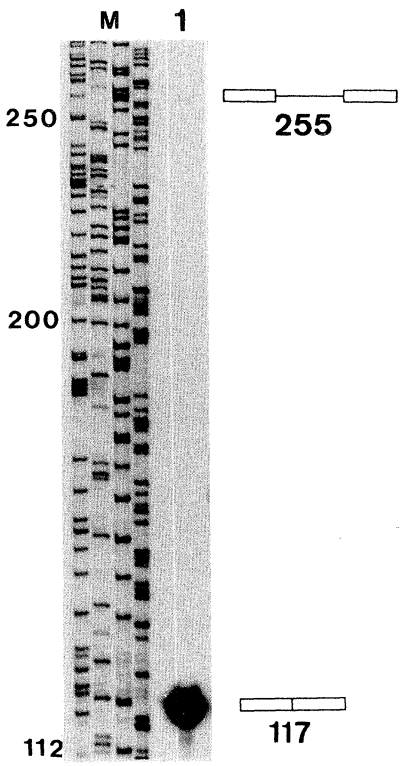


Figure 5. RT–PCR analysis of splicing in transgenic potato. Unspliced and spliced products of the legumin construct are detected in transgenic potato (lane 1) and separated alongside a DNA sequencing marker.

intron constructs in transgenic plants were similar to those obtained in the protoplast transient expression system. For example, in transgenic potato plants, the legumin intron was again accurately and efficiently spliced (figure 5, lane 1). The intron vector and RT–PCR method of analysis is currently being used to analyse 5′ and 3′ splice site selection pathways in plants and the results compared with animal systems, to provide information on branchpoint usage and to analyse the role of AU sequences in intron definition and splicing.

(*c*) *Splicing of an AU-rich antisense intron*

The importance of AU-rich sequences in dicot intron splicing (Goodall & Filipowicz 1989, 1991) has been further confirmed by the analysis of splicing of an effectively non-intron, AU-rich sequence (Simpson & Brown 1993). The AU-rich legumin intron sequence was cloned in reverse orientation in the pDH515 vector such that in transfected protoplasts and transgenic potato plants, the antisense intron sequence, which remains AU-rich, was transcribed as part of the zein transcript (figure 6). RT–PCR analysis of RNA from protoplasts and transgenic plants clearly demonstrated that an efficient splicing event had occurred involving this sequence (Simpson & Brown 1993). Sequencing of the spliced product showed that the AU-rich antisense sequence was spliced by activation of 5′ and 3′ splice site-like sequences, nine nucleotides (nt) upstream and 26 nt downstream respectively of the antisense legumin sequence. Although the role of AU-rich sequences remains to be demonstrated, the splicing of the AU-rich region representing the antisense legumin intron shows that such regions have the intrinsic ability to promote spliceosome formation and splicing. While this finding has implications for plant gene evolution, it demonstrates that the base composition of genes to be introduced into transgenic plants is an important consideration in plant genetic engineering. In particular, genes from heterologous organisms with different codon usage from plants can be more AU-rich than plant gene coding sequences which could promote splicing. The difficulty in demonstrating expression of *Bacillus thuringiensis* toxin in transgenic plants is probably the best known example and the detection of polyadenylated mRNAs shorter than expected may, indeed, suggest cryptic splicing of the toxin mRNA (Murray *et al.* 1991). The strategy of rebuilding the toxin gene to remove splice sites and alter codon usage, which has proved successful for the *Bacillus thuringiensis* toxin gene is now being used before introduction of other heterologous genes.

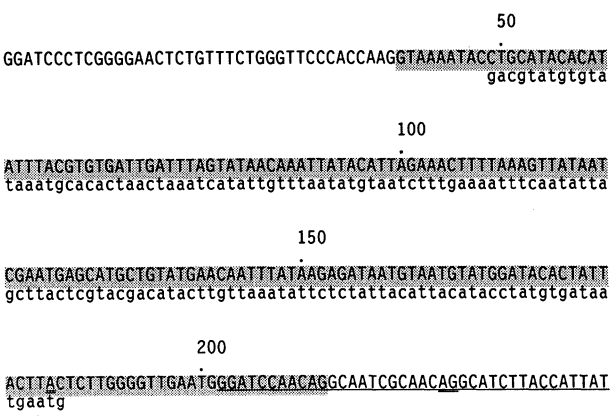


Figure 6. Splicing of an AU-rich antisense intron. The sequence of the transcript containing the antisense legumin intron is given in capitals and the sequences removed by splicing is blocked. The position and sequence of the legumin intron on the other DNA strand is indicated by small letters. Flanking zein sequence is underlined and a second potential 3′ splice site downstream of that chosen is double underlined.

6. PLANT SPLICEOSOMAL PROTEINS

Although many UsnRNP protein genes have been isolated from animal systems and their biochemical functions characterized, little is known about their role in splicing. Transgenic plants may provide a unique opportunity to analyse the function of these proteins in splicing in higher eukaryotes by expression of antisense RNAs or mutated gene sequences. The inclusion, in transformation constructs, of well characterized intron-splicing cassettes will allow assessment of effects on splicing of gene inactivation or protein disruption by antisense or mutation. We have previously isolated the U2snRNP-specific protein, U2B″, from potato (Simpson *et al.* 1991) and have recently isolated a family of RNA helicases (unpublished data). While their biochemical functions (e.g. RNA-binding) can be investigated by *in vitro* systems following production of proteins in *E. coli*, the investigation of function of U2B″ in pre-mRNA splicing and of the RNA helicases or other nuclear proteins in aspects of RNA processing will rely on the production of transgenic plants. Firstly, information on the cellular and sub-cellular location of the proteins may provide clues to potential function, particularly for proteins whose function is unknown, as in the case of the RNA helicases. Of particular interest is whether plant spliceosomal components or proteins, such as U2B″, are located in subnuclear structures which have been observed in animal nuclei (Spector *et al.* 1991). To address this question, we have developed an epitope tag based on a peptide sequence from Potato Leaf Roll Virus coat protein for which a monoclonal antibody is available and epitope tagged U2B″ and RNA helicase constructs are currently being made for transgenic plant production.

One of the most attractive approaches to studying function of plant genes is that of gene inactivation by antisense which has proved to be very effective (reviewed by Watson & Grierson 1992). While transgenic plants containing antisense U2B″ constructs driven by the CaMV 35S promoter are currently being produced, it may be expected that inactivation of proteins which form components of essential cellular processes may be lethal. Therefore, it is important to utilize inducible promoters to be able to control expression of antisense RNA. To this end we are

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

currently using the sucrose-inducible patatin promoter sequence (Jefferson *et al.* 1990) and the *Tet*-repressor system (Gatz *et al.* 1992) to generate inducible antisense U2B″ plants. In this context, if the antisense RNAs prove to inhibit splicing, it may also be feasible to develop *in vivo* complementation systems where endogenous mRNAs are knocked out by an induced antisense RNA and at the same time an inducible protein carrying small mutations expressed to see whether correct and efficient RNA processing is maintained or inhibited. *In vivo* complementation of genetic mutations have been widely used to characterize the function of proteins in yeast and although more complicated in terms of construction and control such an approach may provide a more general method for dissecting functional domains of plant proteins.

Finally, the overexpression of dominant inhibitors in transgenic plants is being used to analyse transcription factors. A similar approach can be adopted to examine the function of essential spliceosomal or RNA processing protein components. For example, the overexpression of mutated or truncated U2B″, RNA helicase or other such proteins may lead to a dominant negative phenotype thereby identifying essential regions in the proteins. We are currently designing transformation constructs to produce dominant negative phenotypes for U2B″ and RNA helicase to examine the efficiency of this approach.

7. CONCLUSIONS

Transient transformation of protoplasts and stable transgenic plants are being used in the study of plant intron splicing and splice site selection. Intron cassettes and methods to specifically detect their splicing and calculate splicing efficiency provide a tool to analyse the effect on splicing of spliceosomal components by incorporating them into, for example, antisense transformation constructs. The cloning of cDNAs for RNA-interacting proteins from plants now allows us to exploit the ease of transgenic plant production in analyses of their function. In many cases where the function of important protein domains can be deduced by analogy to animal or yeast proteins, strategies can be developed to examine the effect of small mutations in *in vivo* plant systems. The use of transgenic plants provides a unique opportunity for the analysis of the extremely complex process of pre-mRNA splicing.

REFERENCES

Abel, S., Kiss, T. & Solymosy, F. 1989 Molecular analysis of eight U1 RNA gene candidates from tomato that could potentially be transcribed into U1 RNA sequence variants differing from each other in similar regions of secondary structure. *Nucl. Acids Res.* **17**, 6319–6337.

Barta, A., Sommergruber, K., Thompson, D., Hartmuth, K., Matzke, M.A. & Matzke, A.J.M. 1986 The expression of a nopaline synthase-human growth hormone chimaeric gene in transformed tobacco and sunflower callus tissue. *Pl. molec. Biol.* **6**, 347–357.

Bennett, M., Pinol-Roma, S., Staknis, D., Dreyfuss, G. & Reed, R. 1992 Differential binding of heterogenous nuclear ribonucleoproteins to mRNA precursors prior to spliceosome assembly in vitro. *Molec. cell. Biol.* **12**, 3165–3175.

Bevan, M. 1984 Binary *Agrobacterium* vectors for plant transformation. *Nucl. Acids Res.* **12**, 8711–8721.

Brown, J.W.S. 1986 A catalogue of splice junction and putative branch point sequences from plant introns. *Nucl. Acids Res.* **14**, 9549–9559.

Brown, J.W.S., Feix, G. & Frendewey, D. 1986 Accurate *in vitro* splicing of two pre-mRNA plant introns in a HeLa cell nuclear extract. *EMBO J.* **5**, 2749–2758.

Callis, J., Fromm, M. & Walbot, V. 1987 Introns increase gene expression in cultured maize cells. *Genes Dev.* **1**, 1183–1200.

Gatz, C., Frohberg, C. & Wendenburg, R. 1992 Stringent repression and homogeneous de-repression by tetracycline of a modified CaMV 35S promoter in intact transgenic tobacco plants. *Plant J.* **2**, 397–404.

Golde, T.E., Estus, S., Usiak, M., Younkin, L.H. & Younkin, S.G. 1990 Expression of β-amyloid protein precursor mRNAs: recognition of a novel alternatively spliced form and quantitation in Alzheimer's disease using PCR. *Neuron* **4**, 253–267.

Goodall, G.J. & Filipowicz, W. 1989 The AU-rich sequences present in the introns of plant nuclear pre-mRNAs are required for splicing. *Cell* **58**, 473–483.

Goodall, G.J. & Filipowicz, W. 1990 The minimum functional length of pre-mRNA introns in monocots and dicots. *Pl. molec. Biol.* **14**, 727–733.

Goodall, G.J. & Filipowicz, W. 1991 Different effects of intron nucleotide composition and secondary structure on pre-mRNA splicing in monocot. and dicot. plants. *EMBO J.* **10**, 2635–2644.

Goodall, G.J., Kiss, T. & Filipowicz, W. 1991 Nuclear RNA splicing and small nuclear RNAs and their genes in higher plants. *Oxf. Surv. Pl. molec. Cell Biol.* **7**, 255–296.

Green, M.R. 1991 Biochemical mechanisms of constitutive and regulated pre-mRNA splicing. *A. Rev. Cell Biol.* **7**, 559–599.

Guerineau, F., Brooks, L. & Mullineaux, P. 1991 Effect of deletion in the cauliflower mosaic virus polyadenylation sequence on the choice of the polyadenylation sites in tobacco protoplasts. *Molec. gen. Genet.* **246**, 141–144.

Guthrie, C. 1991 Messenger RNA splicing in yeast: clues to why the spliceosome is a ribonucleoprotein. *Science, Wash.* **253**, 157–163.

Hanley, B.A. & Schuler, M.A. 1988 Plant intron sequences: evidence for distinct groups of introns. *Nucl. Acids Res.* **16**, 7159–7176.

Hanley, B.A. & Schuler, M.A. 1991 cDNA cloning of U1, U2, U4 and U5snRNA families expressed in pea nuclei. *Nucl. Acids Res.* **19**, 1861–1869.

Hirt, H., Gartner, A. & Heberle-Bors, E. 1992 An alfalfa cDNA encodes a protein with similarity to human snRNP-E. *Nucl. Acids Res.* **20**, 6–13.

Jefferson, R., Goldsbrough, A. & Bevan, M. 1990 Transcriptional regulation of a patatin-1 gene in potato. *Pl. molec. Biol.* **14**, 995–1006.

Keith, B. & Chua, N.H. 1986 Monocot. and dicot. pre-mRNAs are processed with different efficiencies in transgenic tobacco. *EMBO J.* **5**, 2419–2425.

Kenan, O.J., Query, C.C. & Keene, J.D. 1991 RNA recognition: towards identifying determinants of specificity. *TIBS* **16**, 214–220.

Kulesza, H., Simpson, G.G., Waugh, R., Beggs, J.D. & Brown, J.W.S. 1993 Detection of a plant protein analogous to the yeast spliceosomal protein, PRP8. *FEBS Letts.* **318**, 4–6.

Langridge, P., Brown, J.W.S., Pinto-Toro, J.A. & Feix, G.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

1985 Expression of zein genes in *Acetobularis mediterranea*. *Eur. J. Cell Biol.* **39**, 257–264.

Leader, D., Connelly, S., Filipowicz, W., Waugh, R. & Brown, J.W.S. 1993 Differential expression of U5snRNA gene variants in maize (*Zea mays*) protoplasts. *Pl. molec. Biol.* **21**, 133–143.

Lührmann, R., Kastner, B. & Bach, M. 1990 Structure of spliceosomal snRNPs and their role in pre-mRNA splicing. *Biochim. biophys. Acta* **1087**, 265–292.

Madhani, H.D. & Guthrie, C. 1992 A novel base-pairing interaction between U2 and U6snRNAs suggests a mechanism for the catalytic activation of the spliceosome. *Cell* **71**, 803–817.

McCullough, A.J., Lou, H. & Schuler, M.A. 1991 *In vivo* analysis of plant pre-mRNA splicing using an autonomously replicating vector. *Nucl. Acids Res.* **19**, 3001–3009.

Murray, E.E., Rocheleau, T., Eberle, M., Stock, C., Sekar, V. & Adang, M.J. 1991 Analysis of unstable RNA transcripts on insecticide crystal protein genes of *Bacillus thuringiensis* in transgenic plants and electroporated protoplasts. *Pl. molec. Biol.* **16**, 1035–1050.

Palfi, Z., Bach, M., Solymosy, F. & Lührmann, R. 1989 Purification of the major UsnRNPs from broad bean nuclear extracts and characterisation of their protein constituents. *Nucl. Acids Res.* **17**, 1445–1458.

Pietrzak, M., Shillito, R.D., Hohn, T. & Potrykus, I. 1986 Expression in plants of two bacterial antibiotic resistance genes after transformation with a new plant expression vector. *Nucl. Acids Res.* **14**, 5857–5868.

Reddy, A.S.N., Czernik, A.J., An, G. & Poovaiah, B.W. 1992 Cloning of the cDNA for U1 small nuclear ribonucleoprotein particle 70K protein from *Arabidopsis thaliana*. *Biochim. biophys. Acta* **1171**, 88–92.

Rosbash, M. & Séraphin, B. 1991 Who's on first? The U1snRNP-5 splice site interaction in splicing. *TIBS* **16**, 187–190.

Ruby, S.W. & Abelson, J. 1991 Pre-mRNA splicing in yeast. *TIG* **7**, 79–85.

Simpson, C.G. & Brown, J.W.S. 1992 Rapid analysis of plant gene expression by a novel reverse transcriptase-PCR method. *Plant J.* **2**, 835–836.

Simpson, C.G. & Brown, J.W.S. 1993 Efficient splicing of an AU-rich antisense intron sequence. *Pl. molec. Biol.* **21**, 205–211.

Simpson, G.G., Vaux, P., Clark, G., Waugh, R., Beggs, J.D. & Brown, J.W.S. 1991 Evolutionary conservation of the spliceosomal protein U2B″. *Nucl. Acids Res.* **19**, 5213–5217.

Sinibaldi, R. & Mettler, I. 1992 Intron splicing and intron mediated enhanced expression in monocots. *Prog. Nucl. Acids Res. molec. Biol.* **42** (ed. W. E. Cohn & K. Moldave). New York: Academic Press.

Smith, C.W.J., Patton, J.G. & Nadal-Ginard, B. 1989 Alternative splicing in the control of gene expression. *A. Rev. Genet.* **23**, 527–577.

Spector, D.L., Fu, X.-D. & Maniatis, T. 1991 Associations between distinct pre-mRNA splicing components and the cell nucleus. *EMBO J.* **10**, 3467–3481.

Uchimiya, H., Kidou, S., Shimazaki, T., Aotsuka, S., Takamatsu, S., Nishi, R., Hashimoto, H., Matsubayashi, Y., Kidou, N., Umeda, M. & Kato, A. 1992 Random sequencing of cDNA libraries reveals a variety of expressed genes in cultured cells of rice (*Oryza sativa* L.). *Plant J.* **2**, 1005–1009.

van Santen, V.L. & Spritz, R.A. 1987 Splicing of plant pre-mRNAs in animal systems and vice versa. *Gene* **56**, 253–265.

Vankan, P. & Filipowicz, W. 1988 Structure of U2snRNA genes of *Arabidopsis thaliana* and their expression in electroporated plant protoplasts. *EMBO J.* **7**, 791–799.

Vaux, P., Guerineau, F., Waugh, R. & Brown, J.W.S. 1992 Characterisation and expression of U1snRNA genes from potato. *Pl. molec. Biol.* **19**, 959–971.

Waigmann, E. & Barta, A. 1992 Processing of chimeric introns in dicot plants: evidence for a close cooperation between 5′ and 3′ splice sites. *Nucl. Acids Res.* **20**, 75–81.

Wasserman, D.A. & Steitz, J.A. 1992 Interactions of small nuclear RNAs with precursor messenger RNA during in vitro splicing. *Science, Wash.* **257**, 1918–1925.

Watson, C.F. & Grierson, D. 1992 Antisense RNA in plants. In *Transgenic plants: fundamentals and applications* (ed. A. Hiatt), pp. 255–280. New York: Marcel Dekker.

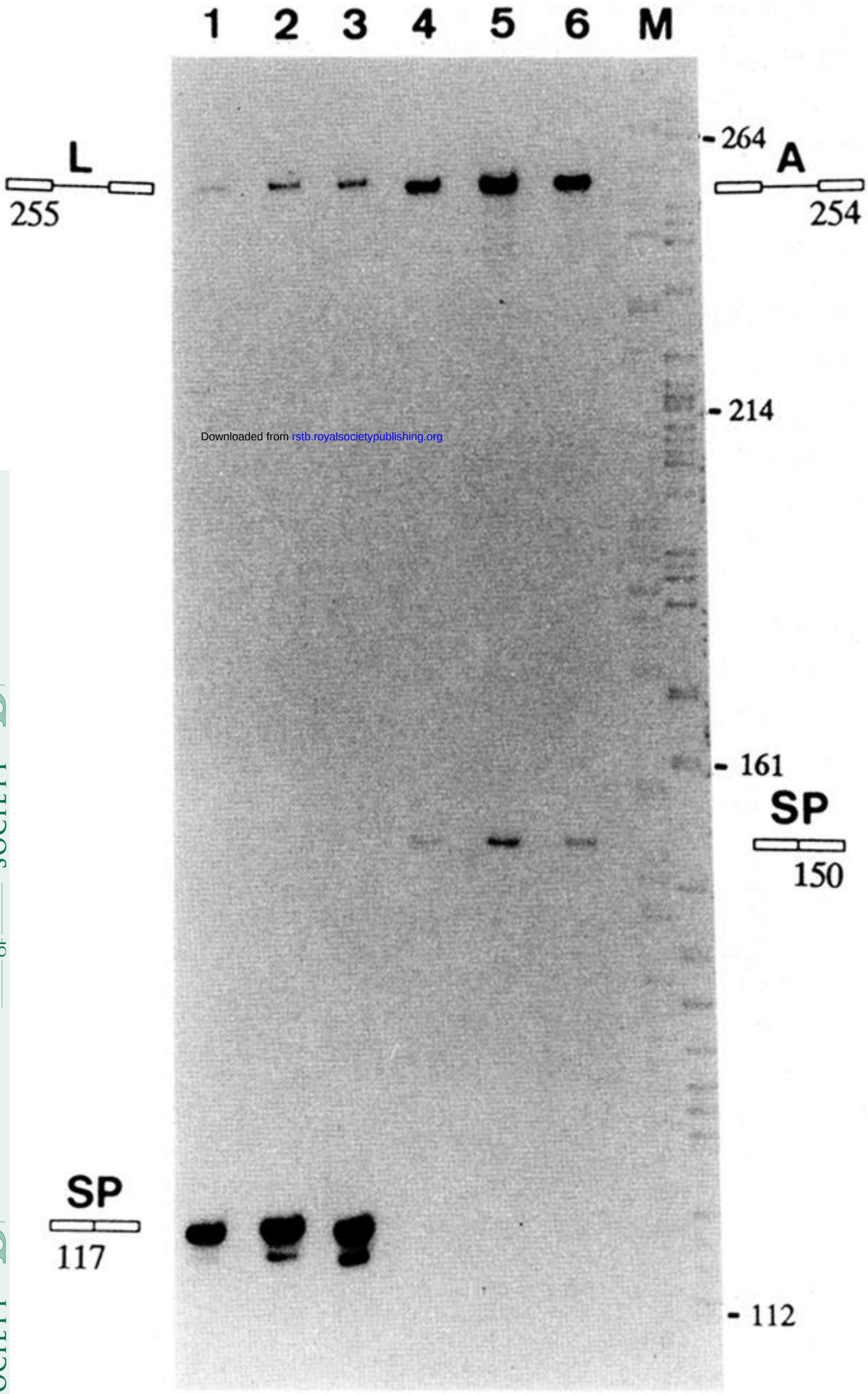
Waugh, R., Clark, G., Vaux, P. & Brown, J.W.S. 1991*a* Sequence and expression of potato U2snRNA gene. *Nucl. Acids Res.* **19**, 249–256.

Waugh, R., Clark, G., Vaux, P. & Brown, J.W.S. 1991*b* Sequence variation and linkage of potato U2snRNA-encoding genes established by PCR. *Gene* **107**, 197–204.

Wernecke, J.M., Chatfield, J.M. & Orgen, W.L. 1989 Alternative mRNA splicing generates the two ribulosebisphosphate carboxylase/oxygenase activase polypeptides in spinach and *Arabidopsis*. *Pl. Cell* **1**, 815–825.

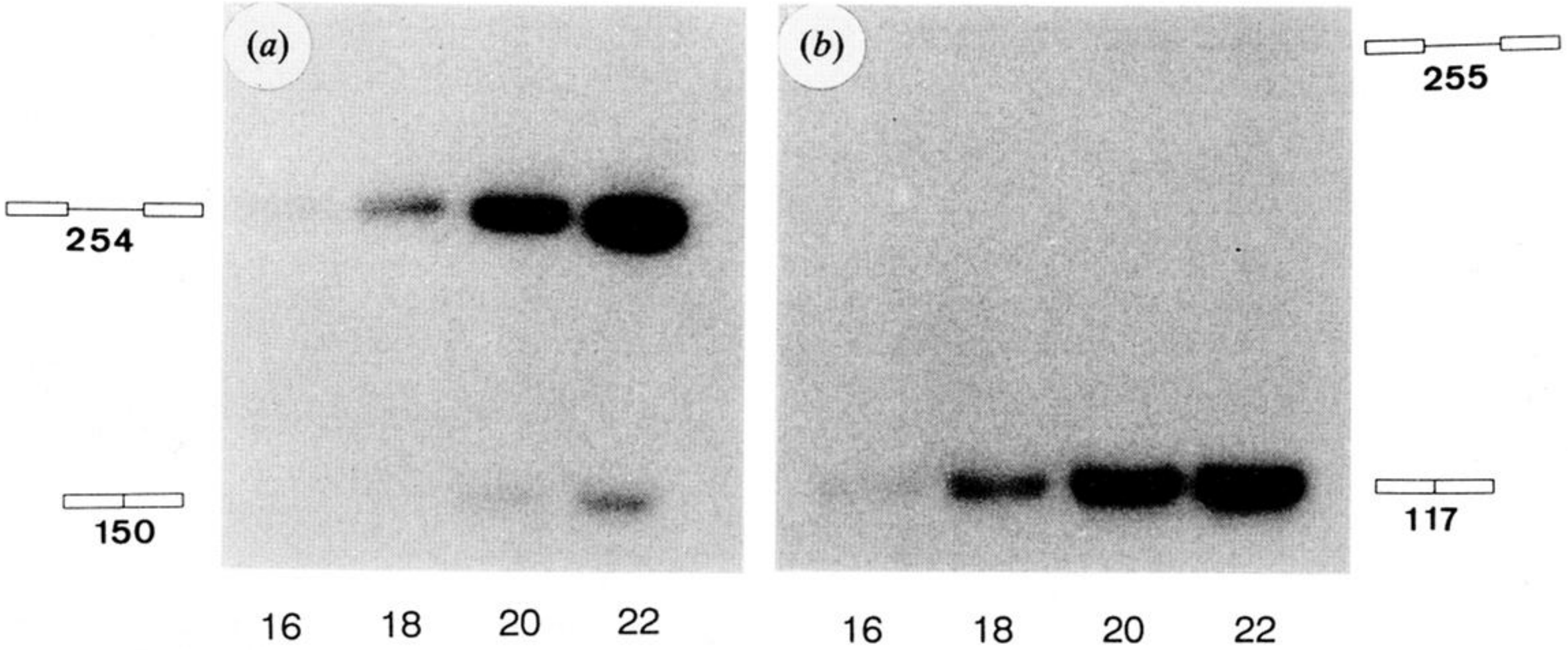
Wiebauer, K., Herrero, J.-J. & Filipowicz, W. 1988 Nuclear pre-mRNA processing in plants: Distinct modes of 3′ splice site selection in plants and animals. *Molec. cell. Biol.* **8**, 2042–2051.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

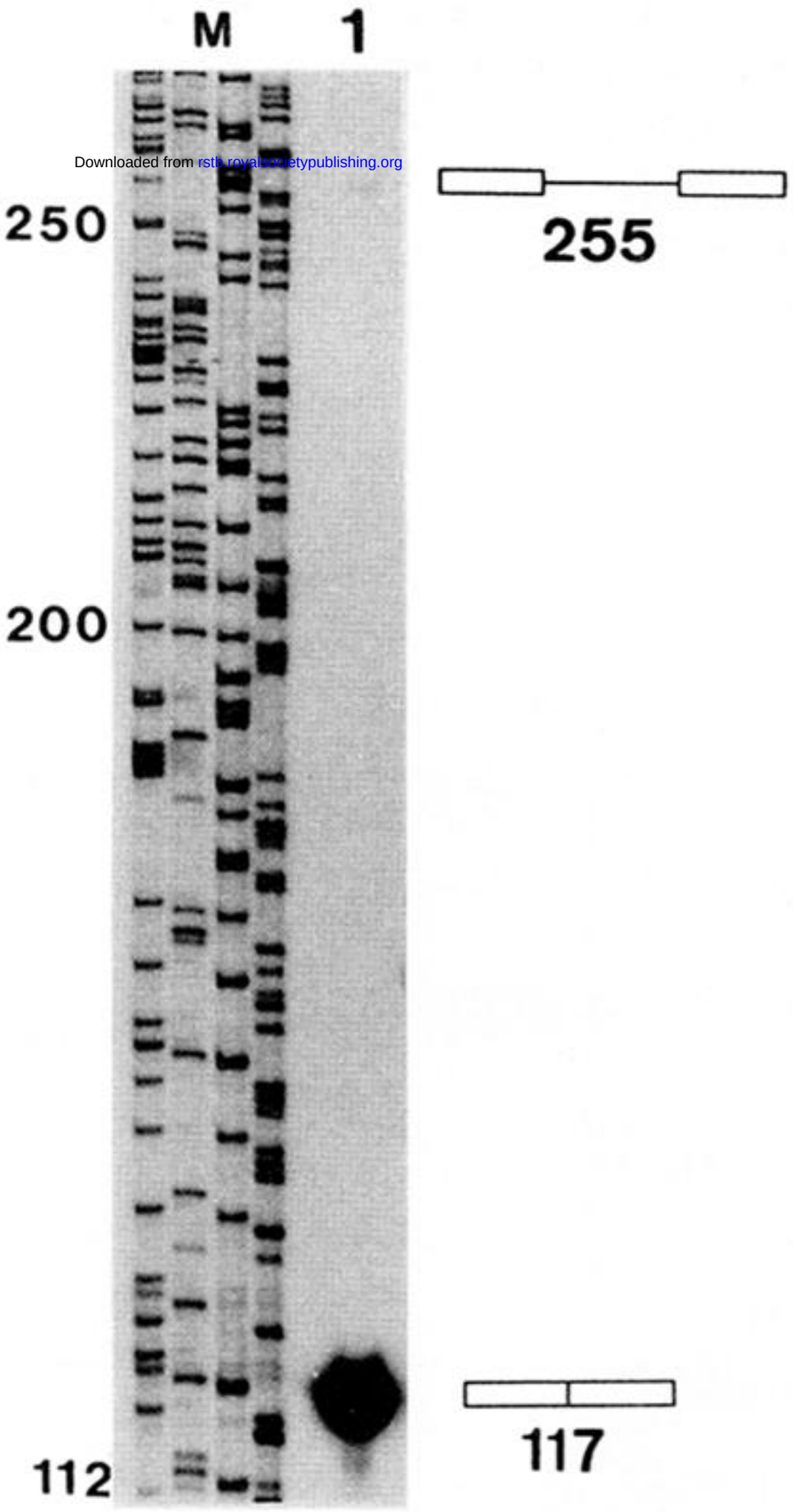


PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

igure 2. RT–PCR analysis of splicing in tobacco proto-lasts. Unspliced (L and A) and spliced products (SP) of hree separate transfections of the legumin intron (lanes 1–3) nd amylase intron (lanes 4–6) are aligned with a DNA equencing marker.



igure 3. Autoradiography of RT–PCR products for quantification. Unspliced and spliced products of the amylase onstruct (*a*) and legumin construct (*b*) from aliquots of the RT–PCR reaction after 16 to 22 cycles were detected by utoradiography following separation on a polyacrylamide gel.



igure 5. RT–PCR analysis of splicing in transgenic potato. nspliced and spliced products of the legumin construct are tected in transgenic potato (lane 1) and separated ongside a DNA sequencing marker.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES