Inefficient and incorrect processing of the Ac transposase transcript in iae1 and wild-type Arabidopsis thaliana

Paul Jarvis¹, François Belzile² and Caroline Dean*  
Department of Molecular Genetics, John Innes Centre,  
Norwich Research Park, Colney, Norwich NR4 7UH, UK

Summary

As part of the analysis of the Arabidopsis mutant iae1-1 (increased Ac excision), quantitative studies of the Ac transposase transcript were conducted. The primary transcript of Ac contains three small introns (introns 1–3; mean size 89 bp) and one larger intron (intron 4; 387 bp). We analysed the splicing of intron 3 and intron 4 in wild-type Arabidopsis and the iae1-1 mutant. Our results demonstrate that the splicing of Ac introns 3 and 4 is inefficient (splicing efficiencies 57 and 30% respectively) compared with that of an intron of an endogenous Arabidopsis gene (PHYB intron 1; splicing efficiency 90%). The poor splicing efficiency of Ac intron 4 was found to correlate with aberrant processing. Steady state levels of total Ac transcript were higher in the iae1-1 mutant than wild-type, but the same aberrant processing occurred. The inefficient processing of Ac in Arabidopsis prompted us to construct an Ac element lacking introns (Ac::cDNA) in an attempt to increase transposition frequencies. Autonomous activity of the Ac::cDNA element was undetectable in Arabidopsis, despite its ability to transpose at high frequency in response to a strong transposase source (35S::transposase) in trans, and the demonstrable autonomy of the same element in tobacco. A number of smaller transcripts were detected in Arabidopsis lines containing Ac::cDNA or Ac. Analysis of these smaller transcripts revealed a high frequency of premature polyadenylation in exon 2 and splicing of cryptic introns.

Introduction

Activator, or Ac, is the autonomous member of a well-characterized family of maize, inverted repeat-type transposable elements initially described by Barbara McClintock in the 1940s [for a review, see Fedoroff (1989)]. The 4565-bp element is structurally unique, possesses 11-bp imperfect inverted repeats at its termini and is flanked by 8-bp target site duplications (Müller-Neumann et al., 1984; Pohlman et al., 1984a,b). A single transcription unit, comprising five exons, four introns and a long (655 bp) untranslated leader sequence, contains a single open reading frame (ORF) of 2421 bp that encodes the 807-amino acid Ac transposase (TPase) protein (Finnegan et al., 1988; Kunze et al., 1987). Ac TPase, the sole element-encoded factor required for transposition, interacts specifically with well defined sequence motifs at the termini of Ac and there, presumably in conjunction with host-encoded factors, catalyses the necessary DNA recombination reactions required for transposition (Coupland et al., 1988, 1989; Kunze and Starlinger, 1989).

Ac has been introduced into a number of heterologous plant hosts with a view to using it as an insertional mutagen in gene-tagging experiments. The activity of Ac was found to be high in both tobacco and tomato, but was very much lower and more variable in Arabidopsis (Dean et al., 1992; Jones et al., 1989, 1991; Yoder, 1990). The reason for the relative inactivity of Ac in Arabidopsis is unclear and may involve DNA methylation, transcript processing or host-encoded transposition repressors. Analyses of the methylation status of Ac in Arabidopsis have so far failed to demonstrate significant levels of DNA methylation (Keller et al., 1992; Lawson et al., 1994), but preliminary evidence for the incorrect processing of Ac transcripts in Arabidopsis was reported by Grevelding et al. (1992). In its native host maize and in tobacco, Ac transcription results in the formation of a single 3.5 kb transcript (Finnegan et al., 1988; Kunze et al., 1987); corresponding to the mature and correctly spliced Ac TPase messenger RNA.

In a previous study, we reported the isolation and initial characterization of a series of Arabidopsis mutants which consistently and heritably displayed very high levels of Ac excision (Jarvis et al., 1996). The mutants, referred to as iae for increased Ac excision, were derived by γ mutagenesis of a line (the iae progenitor) carrying a single copy of the Ac element inserted into the streptomycin phosphotransferase (SPT) excision reporter (Jones et al., 1989). The mutants were identified by their highly variegated phenotypes after growth on streptomycin-containing medium. The mutations were found to fall into two complementation groups, iae1 and iae2, with the iae1 mutants expressing the most extreme phenotypes. The iae mutations were unlinked to the Ac element used to monitor the iae phenotype (SPT::Ac) and segregated in Mendelian fashion. All mutations at the IAE1 locus were found to be recessive, whereas the single mutation at the IAE2 locus was semi-dominant.
In this report, we present the results of detailed analyses of Ac TPase expression in several transgenic Arabidopsis lines (including the most extreme iae1 allele (C4158 or iae1-1) and the iae progenitor). Our primary objective in undertaking these experiments was to assess the role played by Ac TPase transcript abundance and splicing in mediating the markedly elevated levels of Ac excision which we have previously observed in the iae1-1 mutant (Jarvis et al., 1997). However, our analyses also revealed interesting information relating to the general behaviour and expression of Ac in Arabidopsis. Ac transcript processing was found to be highly inefficient and inaccurate in Arabidopsis, resulting in the formation of a large number of aberrant transcripts. We also describe an attempt to improve Ac transposition frequencies in Arabidopsis by the removal of introns from Ac in an effort to alleviate the constraints of inefficient splicing.

Results

Quantitative RT-PCR analysis of Ac TPase transcripts in Arabidopsis

Northern blot analysis of Ac transcripts was initially attempted, but the very low abundance of Ac transcripts in Arabidopsis was found to preclude accurate band resolution (data not shown). We therefore adopted a quantitative reverse transcriptase–polymerase chain reaction (RT–PCR) approach to analyse Ac transcripts in the iae progenitor and iae1-1 mutant (Golde et al., 1990); both lines carry the same T-DNA insertion which harbours a single copy of the Ac element inserted into the streptomycin phosphotransferase (SPT) excision reporter. Quantitative results were obtained by determining the amount of product at several points during the exponential phase of amplification (16–22 cycles), and then using regression analysis to calculate the amount of each template present at the beginning of the reaction (i.e., at zero PCR cycles). This technique is quantitative since it takes into account the differences in amplification efficiency of different PCR products and reactions.

By utilizing PCR primers flanking introns within the Ac gene, we were able to assess splicing efficiencies as well as steady-state transcript levels using the RT–PCR technique. The Ac TPase gene comprises five exons separated by three small introns (introns 1, 2 and 3; 107, 71 and 89 bp, respectively) and one large intron (intron 4; 387 bp). A detailed analysis of the splicing of two introns, one small intron (intron 3) and one large intron (intron 4) was undertaken. PCR primers were, therefore, synthesized flanking these two introns (primers INT3-F and INT3-R for intron 3, and INT4-F and INT4-R for intron 4) and RT–PCR analysis was carried out individually for each intron. The results of these RT–PCR experiments are shown in Figure 1(a) and 1(b). An additional RT–PCR experiment was carried out using the forward primer used in the analysis of intron 3 (INT3-F) and the reverse primer used in the analysis of intron 4 (INT4-R) as the PCR primer set; the result of this experiment is shown in Figure 1(c). As a control to normalize cDNA concentrations (the same cDNA preparations were used in each experiment), RT–PCR analysis of the splicing of an endogenous gene (PHYB; Reed et al., 1993) was also carried out; the result of this experiment is shown in Figure 1(d). The quantitative data derived from all of these experiments (Figures 1a–d are shown in Table 1). All amplified products observed in these RT–PCR experiments correspond to polyadenylated transcripts since oligo (dT) was used to prime the reverse transcription reactions and PCR in the absence of reverse transcription produced no amplification.

Conclusions relating to the analysis of the iae1-1 mutant which can be drawn from the data shown in Table 1, are as follows: (i) there is significantly more Ac TPase transcript in the iae1-1 mutant than in the iae progenitor (this was determined to be an increase of between six and 11-fold); (ii) the majority of the additional Ac transcript in the iae1-1 background is in the unspliced form (there is 13–14 times more unspliced message in the iae1-1 mutant, but only 1.6–3.0 times more spliced message); and (iii) the splicing efficiency of the Ac primary transcript is not improved in the iae1-1 background (the splicing efficiency of Ac intron 3 in the progenitor and iae1-1 backgrounds was determined to be 57 and 14%, respectively).

These results yield a number of conclusions that relate to the expression and general behaviour of Ac in Arabidopsis. The first of these is that the introns of the Ac transposable element are spliced very inefficiently by comparison with an intron of an endogenous Arabidopsis gene. The splicing efficiency of PHYB intron 1 was found to be 90% in the progenitor line, whereas the splicing efficiencies of Ac introns 3 and 4 were determined to be 57 and 30%, respectively in the same line. The second is that the splicing efficiency of Ac intron 4 (30% in the progenitor) is substantially less than that of Ac intron 3 (57% in the progenitor). Furthermore, the splicing efficiency of both introns together (Figure 1c; 30% in the progenitor) was found to be the same as that observed for intron 4 alone, suggesting that the rate limiting factor for Ac transcript processing is the splicing of intron 4.

The analysis of Ac intron 4 splicing (Figure 1b) revealed the presence of four unexpected PCR products. These additional PCR products (estimated to be 550, 530, 450 and 420 bp in length) were intermediate in size between those corresponding to the immature, unspliced transcript (604 bp) and the correctly spliced transcript (217 bp). The analysis of the splicing of introns 3 and 4 together (Figure 1e) revealed the presence of five major, additional PCR products (estimated to be 740, 685, 665, 580 and
Aberrant Ac transcripts in Arabidopsis

550 bp in length) and two others of very much lower abundance (estimated to be 480 and 440 bp in length). All of these additional PCR products are intermediate in size between those corresponding to the unspliced (826 bp) and mature (350 bp) transcripts. We interpret the 740 bp product to represent the correct splicing of intron 3 only (826 – 89 = 737 = 740), and the other four major, additional bands to correspond to the splicing of intron 3 in combination with the same aberrant processing events observed in Figure 1(b) (826 – 604 – 89(=)133; (685, 665, 580 or 550) – 133=(552, 532, 447 or 417)=(550, 530, 450, 420)). That PCR products equivalent to the 480 and 440 bp products observed in Figure 1(c) were not observed in Figure 1(a) or (b) indicates that the aberrant transcripts to which they correspond lack the region of exon 4 to which the INT3-R and INT4-F primers anneal.

The Ac::CiaI/EcoRV and Ac::cDNA constructs

Given that the Ac primary transcript is processed with such poor efficiency and accuracy in Arabidopsis, we reasoned that an Ac element lacking introns within its TPase coding sequence might display improved TPase expression and thereby higher levels of transposition. The construction of such an Ac element was achieved by introducing two restriction sites into the wild-type element and a TPase cDNA sequence (a Clal site near the first ATG translation initiation point, and an EcoRV site near the translation termination codon) enabling the simple removal of the genomic TPase coding sequence and its replacement with the cDNA version. The doubly mutagenized wild-type element (Ac::Clal/EcoRV) was used as a control in all experiments with the cDNA-containing element (Ac::cDNA) in case the small sequence alterations, necessary to produce the required restriction sites, affected element mobility. The modified Ac elements were introduced into a streptomycin phosphotransferase (SPT) fusion, previously used to monitor element activity (Dean et al. 1992, Jones et al., 1989). The functionality of the Ac cDNA employed in this study had previously been demonstrated in tobacco (Scofield et al., 1993). The structures of both constructs are illustrated in Figure 2.

Figure 1. Quantitative RT-PCR analysis of Ac TPase transcripts in Arabidopsis.
The expression of Ac TPase was investigated in two Arabidopsis lines (iae progenitor and iae-1 mutant) using quantitative RT-PCR. Determination of the amount of amplified product at several points during the exponential phase of amplification (16-22 cycles) enabled the amount of template present prior to amplification to be calculated using regression analysis (Table 1). PCR primers flanking the following introns were employed: (a) Ac intron 3; (b) Ac intron 4; (c) Ac introns 3 and 4; and (d) PHYB intron 1. PCR products corresponding to unspliced (U) and spliced (S) transcripts are indicated in each case. Estimated sizes of aberrant transcripts are also shown.
The transformants carrying Ac::Clal/EcoRV and Ac::cDNA constructs in Arabidopsis

Arabidopsis transformants, carrying either the Ac::Clal/EcoRV or Ac::cDNA constructs, were obtained. Analysis of the segregation of kanamycin resistance (conferrered by the neomycin phosphotransferase II [NPT II] gene) in the T2 and T3 generations enabled the selection of three homozygous, single locus transformants of each genotype for subsequent analyses. Molecular analysis of the T-DNA insertions showed that (i) two of the transformants carrying Ac::Clal/EcoRV and two carrying Ac::cDNA contained simple, single-copy and complete insertions of the T-DNA; (ii) one of the single locus transformants carrying Ac::Clal/EcoRV also carried a transposed Ac element; (iii) a third transformant carrying Ac::Clal/EcoRV had a complex T-DNA insertion but did harbour at least one complete copy of the correct Ac element and (iv) one of the transformants carrying Ac::cDNA had two copies of the T-DNA left border but no right border or Ac-hybridizing sequences. This transformant was therefore not considered in subsequent analyses.

Approximately 200 progeny from each of the three homozygous lines carrying Ac::Clal/EcoRV and from the two carrying Ac::cDNA were germinated and grown on streptomycin-containing medium in order to assess Ac excision frequencies. Variegation was scored after 2 weeks’ growth, and the results are shown in Table 2. These data demonstrate that the introduction of the Clal and EcoRV restriction sites into the Ac sequence did not alter the ability of the element to excise since the transformants carrying Ac::Clal/EcoRV all displayed expected levels of Ac activity (mean frequency of variegated individuals=8.1%). The results also show that the Ac::cDNA element is completely inactive in Arabidopsis (no variegation was observed). The reason for the inactivity of Ac::cDNA in Arabidopsis was investigated by testing (i) its ability to trans-activate a Ds element, and (ii) its ability to be trans-activated by a strong TPase source (35S::TPase).

The three transformants carrying Ac::cDNA (and, as controls, the three transformants carrying Ac::Clal/EcoRV) were crossed to a transgenic line [C12b; Bancroft and Dean (1993)] carrying a Ds element inserted into the SPT excision reporter (SPT::Ds). The Ds element present in the C12b line is highly responsive to TPase expression in trans and has the ability to excise at high frequency (Bancroft and Dean, 1993). Variegation frequencies were scored in the F1 and F2 generations, and the results are shown in Table 2. Expected levels of variegation were observed in both F1 and F2 populations derived from crosses between C12b and the transformants carrying Ac::cDNA. This demonstrates that the Ac::cDNA element is unable to trans-activate the C12b Ds element, most likely due to an inability to encode active TPase.

The transformants carrying Ac::cDNA and Ac::Clal/EcoRV were also crossed to a transgenic line carrying the genomic Ac TPase coding sequence fused to the strong CaMV 35S
The Ac::Clal/EcoRV and Ac::cDNA constructs.
Schematic representations of the T-DNA constructs used to assess the activity of the Ac::Clal/EcoRV and Ac::cDNA elements. Also indicated are the LB, RB and Ac probes and the relevant Hindlll and Hpal restriction sites used in the analysis of the T-DNA insertions present in the Arabidopsis transformants.

Table 2. Genetic analysis of the behaviour of Ac::Clal/EcoRV and Ac::cDNA in Arabidopsis

<table>
<thead>
<tr>
<th>Construct</th>
<th>Frequency of variegated seedlings (%)</th>
<th>Frequency of germinal revertants (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ac::Clal/EcoRV (Arabidopsis homozygotes)</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>Ac::cDNA (Arabidopsis homozygotes)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Ac::Clal/EcoRV × SPT::Ds F1</td>
<td>15</td>
<td>0</td>
</tr>
<tr>
<td>Ac::Clal/EcoRV × SPT::Ds F2</td>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td>Ac::cDNA × SPT::Ds F1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Ac::cDNA × SPT::Ds F2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Ac::Clal/EcoRV × 3SS::TPase F1</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>Ac::Clal/EcoRV × 3SS::TPase F2</td>
<td>20</td>
<td>12</td>
</tr>
<tr>
<td>Ac::cDNA-1 × 3SS::TPase F1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Ac::cDNA-1 × 3SS::TPase F2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Ac::cDNA-8 × 3SS::TPase F1</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>Ac::cDNA-8 × 3SS::TPase F2</td>
<td>14</td>
<td>41</td>
</tr>
<tr>
<td>Ac::Clal/EcoRV (tobacco homozygotes)</td>
<td>79</td>
<td>13</td>
</tr>
<tr>
<td>Ac::cDNA (tobacco homozygotes)</td>
<td>53</td>
<td>13</td>
</tr>
</tbody>
</table>

The data shown are the somatic (as assayed by variegation on streptomycin-containing medium) and germinal (as assayed by full resistance on streptomycin-containing medium) excision frequencies of the Ac::Clal/EcoRV and Ac::cDNA elements in homozygous Arabidopsis transformants and in F1 and F2 progeny after crosses to transformants carrying SPT::Ds (C12b) or 3SS::TPase. Excision frequencies of the same elements in tobacco transformants are also shown. In most cases, frequencies were calculated by averaging data from individual transformants carrying the same construct. Data for the F1 and F2 progeny from crosses between two Arabidopsis transformants carrying Ac::cDNA and 3SS::TPase were very different, so both have been included.

The line had been shown to drive high frequencies of Ds excision in both F1 and F2 generations (Swinburne et al., 1992). Variegation was therefore scored in F1 and F2 generations after 2 weeks' growth on streptomycin-containing medium and the results obtained are shown in Table 2. In the F1 population derived from crosses between 3SS::TPase and the transformants carrying Ac::Clal/EcoRV, very high levels of variegation were observed as expected (frequency of variegated individuals=100%); in the F2 generation of the same cross, high levels of variegation and germinal reversion were observed (mean frequency of variegated individuals=23.1%; mean frequency of germinal revertants=13.5%). In F1 and F2 populations derived from crosses between 3SS::TPase and one of the transformants carrying Ac::cDNA (BL1), no variegation or germinal reversion was observed; this demonstrates that the Ac::cDNA element present in this transformant is defective with...
respect to the cis-acting requirements for transposition. However, very high levels of Ac::cDNA excision were observed in both F_1 and F_2 generations derived from crosses between 35S::TPase and transformant B8, carrying Ac::cDNA (frequency of variegated F_1 individuals=100%; frequency of variegated F_2 individuals=14.3%; frequency of F_2 germinal revertants=41.2%). The degree of variegation in these F_1 populations was markedly higher than in those derived from crosses between 35S::TPase and the transformants carrying Ac::Clal/EcoRV; the significantly higher frequencies of germinal reversion in the F_2 generation are presumably a manifestation of the observed higher levels of somatic activity. These data demonstrate that the Ac::cDNA is fully capable of excision from the genome.

**Behaviour of the Ac::Clal/EcoRV and Ac::cDNA constructs in tobacco**

In order to investigate further the inactivity of the Ac::cDNA element in *Arabidopsis*, its activity in tobacco was analysed. As a control, the behaviour of Ac::Clal/EcoRV in tobacco was also investigated. Tobacco transformants carrying both constructs were obtained. The primary transformants (five of each line) were allowed to self-pollinate, and the T_2 progeny of each transformant were scored on medium containing either kanamycin or streptomycin. Four of the five transformants carrying each construct contained single T-DNA loci, whereas the fifth in each case contained two loci.

The main conclusion to be drawn from the results presented in Table 2 is that the Ac::cDNA element is not inherently defective. The data clearly demonstrate substantial Ac::cDNA activity within the tobacco transformants. The behaviour of Ac::cDNA, however, was different from that of Ac::Clal/EcoRV. Firstly, Ac::cDNA was less active than Ac::Clal/EcoRV (overall frequencies of variegated T_2 individuals were 53% and 79%, respectively). Secondly, the activity of Ac::cDNA was significantly more variable, between different tobacco transformants, than Ac::Clal/EcoRV.

**RT-PCR analysis of Ac TPase transcripts in the Arabidopsis transformants carrying Ac::Clal/EcoRV and Ac::cDNA**

In an attempt to elucidate the reason for the non-autonomy of the Ac::cDNA in *Arabidopsis*, we analysed steady-state Ac transcript levels using RT-PCR in the transformants carrying Ac::Clal/EcoRV and Ac::cDNA, the iae progenitor and iae1-1. Primers flanking all four Ac introns (INT1-F and INT4-R) were used in this analysis in order to obtain as much information as possible about the nature of Ac transcripts present in the various transgenic lines. Amplification was carried out for 20 cycles (maximum extent of exponential amplification), and the reactions were then electrophoresed through agarose and blotted. The blot was probed with an internal fragment of Ac, and the result obtained is shown in Figure 3(a). In order to normalize cDNA concentrations, RT-PCR was carried out in similar
fashion using primers and a probe specific for an endogenous Arabidopsis gene (FCA; Macknight and Dean, unpublished data). The result of this experiment is shown in Figure 3(b). Since different PCR primers were used in these experiments and those shown in Figure 1, it is difficult to compare the transcripts directly.

Figure 3(a) demonstrates that each of the seven Ac-containing transgenic lines expresses multiple and diverse Ac transcripts. Predicted PCR fragment sizes for this analysis were 3.21 kb for the primary Ac transcript (all four introns retained) and 2.56 kb for the mature, fully-spliced Ac transcript. Whereas the 2.56-kb product was observed, with varying abundance, in all seven Ac-containing transformants, only two PCR products larger in size than the 2.56-kb product were observed and the 3.21-kb product was undetectable (with the possible exception of a very faint band in the iae-1-1 track). The majority of PCR products corresponded to aberrant Ac transcripts and ranged in size from 2.94 to 0.63 kb.

In agreement with the data in Figure 1, the steady-state abundance of Ac TPase transcript in the iae-1-1 background was found to be higher than in the iae progenitor. The bulk of the additional Ac message present in the iae-1-1 mutant exists in the form of five alternative transcripts of abnormally low molecular weight (corresponding PCR products: 1.28, 1.22, 1.05, 0.96 and 0.91 kb); the abundance of mature 3.5-kb transcript is largely unaltered in the mutant. These alternative transcripts are not unique to the iae-1-1 background since they are also present in the progenitor at much lower abundance.

The pattern observed for transformants carrying Ac::Clal/EcoRV is very similar to that observed for the progenitor and iae-1-1, but that observed for transformants carrying Ac::cDNA is clearly distinct (Figure 3a). Seven transcripts are unique to Ac::cDNA transformants (corresponding PCR products: 1.49, 1.43, 0.86, 0.81, 0.75, 0.67 and 0.63 kb). The mature Ac transcript (corresponding PCR product: 2.56 kb) is present in transformants carrying Ac::cDNA but at very much reduced (approximately 10-fold) abundance compared with transformants carrying Ac::Clal/EcoRV.

In order to investigate further the nature of the smaller transcripts generated in both Ac::Clal/EcoRV and Ac::cDNA-containing lines, seven RT–PCR products were cloned and sequenced. Six of the seven clones analysed were the result of premature termination and polyadenylation at sites between nucleotides 1885 and 1913 within exon 2 (Figure 4). Four of these were transcripts from Ac::Clal/EcoRV and two from Ac::cDNA. The seventh clone analysed (a transcript from a transformant carrying a Ac::Clal/EcoRV element) showed correct splicing of introns 1, 2, 3 and 4 but also carried a 505 bp deletion within exon 3 (bp 2737–3241 inclusive) which we interpret to represent a cryptic intron. The first two base pairs of this intron are GT and the last two are AG, both of which conform to the normal splice site consensus.

**Discussion**

The splicing of Ac introns in Arabidopsis

We have conducted quantitative RT–PCR analyses of Ac TPase expression in the iae-1-1 mutant and in the transgenic progenitor line from which the mutant was derived (Figures 1 and 3 and Table 1). Our results demonstrate that steady-state levels of total Ac transcript are higher in the iae-1-1 mutant than in the progenitor (13– to 14-fold), but that the majority of additional Ac message present in the mutant exists in the form of unspliced or alternatively processed transcript. The increase in abundance of the mature transcript was only small (1.6- to 3.0-fold) and unlikely to account for the previously observed 550-fold increase in Ac excision frequency (Jarvis et al., 1997).

We suggest that factors other than Ac TPase transcript abundance and processing are responsible for the iae-1-1 phenotype. The elevated abundance of Ac transcripts (which are unlikely to encode functional TPase protein) in the iae-1-1 mutant may be an indirect result of the mutation and a consequence of the elevated excision frequencies.

Our analyses also revealed general information relating to the behaviour and expression of Ac in wild-type Arabidopsis. Firstly, the splicing efficiency of Ac introns was found to be significantly less than the splicing efficiency of an endogenous PHYB intron; secondly, the splicing efficiency of Ac intron 4 was significantly less than that of Ac intron 3; and finally, the particularly low splicing efficiency of Ac was found to correlate with its alternative processing. The poor splicing of Ac introns in Arabidopsis cannot be explained easily in terms of the splice junction sequences since these are highly conserved, not only between different plant species, but also across the plant and animal kingdoms (Brown, 1988). However, it has previously been demonstrated that the requirements for efficient intron processing differ between monocotyledinous and dicotyledinous species (Goodall and Filipowicz, 1991; Keith and Chua, 1986). AU-rich sequences are a characteristic feature of dicotyledonous introns (minimum AU=60%) and are an essential requirement for their efficient splicing; the presence of GC-rich introns in monocotyledonous plants demonstrates that AU-rich sequences are not essential (Goodall and Filipowicz, 1989, 1991). The AU nucleotide content of the introns and exons of Ac and of the endogenous Arabidopsis gene PHYB are shown in Table 3. The poor splicing efficiency of Ac intron 3 is difficult to explain in terms of its AU content (76.40%). Exon sequences adjacent to Ac intron 3 are also AU-rich (exon 3: 63.24% AU; exon 4: 59.78% AU), and it is therefore possible that the high AU content of Ac exons 3 and 4 masks the
Figure 4. Sequence analysis of seven short, aberrantly processed Ac transcripts.

(a) Polyadenylation sites within exon 2 are shown for six prematurely terminated transcripts underneath the sequence of Ac from bp 1880 to 1930. The clones were derived from either Ac::Clal/EcoRV or Ac::cDNA transformants or from the iael-1 line (C4158).
(b) The site of a 505 bp cryptic intron within exon 3 in a transcript from a transformant carrying Ac::Clal/EcoRV. Numbers in (a) and (b) denote base pairs from the 5' end of Ac and run consecutively from 1 to 4565.

Table 3. Nucleotide compositions of the introns and exons of Ac TPase and of the Arabidopsis PHYB gene

<table>
<thead>
<tr>
<th></th>
<th>Ac Exon</th>
<th>Ac Intron</th>
<th>PHYB Exon</th>
<th>PHYB Intron</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>52.42 (124)</td>
<td>71.03 (107)</td>
<td>53.88 (2164)</td>
<td>62.50 (88)</td>
</tr>
<tr>
<td>2</td>
<td>59.35 (1348)</td>
<td>74.65 (71)</td>
<td>54.08 (808)</td>
<td>70.00 (90)</td>
</tr>
<tr>
<td>3</td>
<td>63.24 (797)</td>
<td>76.40 (89)</td>
<td>55.67 (291)</td>
<td>67.02 (376)</td>
</tr>
<tr>
<td>4</td>
<td>59.78 (92)</td>
<td>58.66 (387)</td>
<td>52.57 (253)</td>
<td>–</td>
</tr>
<tr>
<td>5</td>
<td>61.67 (60)</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

The composition of each exon and intron, expressed as the percentage of A+U nucleotides per total number of nucleotides, is given. The total number of nucleotides in each exon or intron is given in parentheses. These data are an adaptation of those presented by Pohlman et al. (1984a,b) and Reed et al. (1993).

The presence of Ac intron 3. Ac intron 4 is approximately four times larger (387 bp) than introns 1, 2 and 3 and has a substantially lower proportion of A and U nucleotides in its sequence (58.66% AU). This AU content could make it essentially indistinguishable from the AU-rich flanking exons.

Analysis of Ac transcripts in Arabidopsis transformants carrying Ac::Clal/EcoRV and Ac::cDNA

The RT–PCR analysis of Ac transcripts in the transformants carrying Ac::Clal/EcoRV and Ac::cDNA elements, iae progenitor and iael-1 lines, using PCR primers flanking the sites of all four introns, revealed the presence of a large number of transcripts in all seven Ac-expressing lines investigated. The majority of these transcripts (16 out of a...
total of 18 distinct, aberrant transcripts) were smaller in size than the mature and correctly processed message. Analysis of seven of these aberrantly spliced transcripts showed that a major mis-processing step is premature termination and polyadenylation at closely linked sites within exon 2. However, one of the clones showed some evidence of the splicing of a 505 bp cryptic intron within exon 3. Thus, the range of smaller transcripts comes from a number of mis-processing events. In the absence of a 5' splice site, polyadenylation within AU-rich sequences has been shown to be the default pathway (Luehrsen and Walbot, 1994). The even smaller transcripts from the intronless Ac element construct are therefore likely to result from increased mis-processing due to the absence of splice sites.

This analysis of Ac transposase transcript types and levels has provided much information to account for the behaviour of Ac in Arabidopsis. It has not provided clear evidence for the nature of the role performed by the IAE1 locus. Further work will be required to elucidate the factors that act to limit Ac activity in Arabidopsis which are alleviated in the iae1-1 mutant.

Experimental procedures

Construction of Ac::ClaI/EcoRV and Ac::cDNA

Oligo mutagenesis using the primers GCTACGACTCCATTCCATCGATGACGCCTC (ClaI) and CTCTCCATGAGCAAGATCTTATTATGGTTG (EcoRV) was used to introduce a ClaI site near the first translation initiation codon and an EcoRV site at the translation stop codon into an Ac TPase cDNA clone (Scofield et al., 1993); the resultant plasmid was termed SLJ321B5. The same oligonucleotides were used to introduce ClaI and EcoRV sites into an Ac genomic clone (Ac::ClaI/EcoRV; the plasmid was termed SLJ322B1). The ClaI/EcoRV cDNA fragment of SLJ321B5 was used to replace the genomic ClaI/EcoRV fragment of SLJ322B1 to give an Ac element with intact 5' and 3' regions but carrying no introns (Ac::cDNA; the plasmid was termed SLJ3241). The SacII/SacI fragments of SLJ3241 and SLJ322B1 were each cloned into the 5' untranslated leader of the 35S–SPT fusion carried by the binary vector, CL0111 (Dean et al., 1992).

Plant materials and tissue culture procedures

All Arabidopsis thaliana lines used in this study were of the ecotype Landsberg erecta. Two previously described transgenic Arabidopsis lines were used: the iae progenitor (B1798) and the iae1-1 mutant (C4158), both of which contain the 0194 SPT::Ac reporter fusion (Dean et al., 1992; Jarvis et al., 1997). Arabidopsis tissue culture and in vitro growth procedures, including transformation and the scoring of the streptomycin and kanamycin resistance phenotypes, were carried out as described by Valvekens et al. (1988), Dean et al. (1992). Tobacco (Nicotiana tabacum cv. Petite Havana) transformation was performed according to the protocol described by Horsch et al. (1985) using the same Agrobacterium tumefaciens strains (C58C1; Deblaere et al., 1985) used to transform Arabidopsis. Tobacco kanamycin and streptomycin resistant phenotypes were scored as described by Horsch et al. (1985) and Scofield et al. (1993).

Southern blotting

A genomic DNA preparation was carried out as described by Dean et al. (1992). Southern blotting was performed according to the Hybond N protocol (Amersham). The LB probe used was a 1.0-kb ClaI/SphI fragment carrying the 1' and 2' promoters and the 5' end of the NPT II gene; the RB probe was a 0.4-kb SspI/XhoI fragment carrying the CaMV 35S promoter and the T-DNA RB sequences; Ac probes used were the entire 4.6-kb element and the 1.6 kb internal HindIII fragment (Figure 2; Dean et al., 1992).

RT-PCR analysis of Ac transcripts

Total RNA was extracted from plant tissue according to the procedure described by Logemann et al. (1987) and treated twice with DNasel (FPLCpure; Pharmacia) in order to remove any contaminating DNA. Total RNA samples (10 μg) were mixed with 20 pmol oligo (dT) primer in a total volume of 10 μl, denatured at 70°C for 10 min and then cooled rapidly on ice. Reverse transcription was carried out at 37°C for 60 min in 50 mM Tris–HCl, pH 8.3, 75 mM KCl, 3 mM MgCl2, 10 mM DTT, 0.5 mM dATP, 0.5 mM dCTP, 0.5 mM dGTP and 0.5 mM dTTP using 200 units superscript reverse transcriptase ( Gibco BRL) and in the presence of 35 units RNase inhibitor (RNAGuard; Pharmacia). Aliquots (1 μl) of the cDNA thus produced were subjected to PCR amplification. The PCR primers used were as follows: INT3-F (GCTTCGTTCCAGTGCTGTT), INT3-R (AGTCTAGACCTCGAGATCACC), INT4-F (ATGGATGGTGATCGTGCCAGCT), INT4-R (CACCTGTCACATGCA) and INT1-F (TCCAGGGCTCTGCTCCACTTCGCCG) for the analysis of Ac transcripts: F12 and R14 for the analysis of PHYB transcripts (Bradley et al., 1995); and cDNA1-F and 3' A-R for FCA (Macknight and Dean, unpublished data).

The quantitative RT-PCR protocol employed (Figure 1) was an adaptation of that described by Golde et al. (1990). Forward (F) and reverse (R) primers were end-labelled using T4 polynucleotide kinase and [γ-32P]dATP prior to PCR amplification. Unincorporated nucleotides were removed by ethanol precipitation in the presence of ammonium acetate (2.0 M prior to addition of ethanol). PCR amplification conditions were as follows: 1 min at 94°C, 2 min at 55°C and 3 min at 72°C for up to 22 cycles. Aliquots (5 μl) were removed from the PCR (total volume 50 μl) after increasing numbers of cycles during the exponential phase of amplification (16, 17, 18, 19, 20, 21 and 22 cycles) and then electrophoresed through denaturing polyacrylamide gels. Gels were transferred to Whatman 3M paper and dried; bands were visualized and quantified using a phosphorimager (Molecular Dynamics). Using regression analysis, the amount of each template present at the beginning of the PCR reaction (i.e. at zero cycles) was calculated. This RT-PCR technique is quantitative since it takes into account the differences in amplification efficiency of different PCR products and reactions.

For the analysis shown in Figure 5, PCR amplification conditions were as follows: 1 min at 94°C, 2 min at 58°C and 5 min at 72°C for 20 cycles (Ac); and 1 min at 94°C, 2 min at 60°C and 3 min at 72°C for 20 cycles (FCA). Amplification was preceded by denaturation at 94°C for 5 min. Aliquots (20 μl) of each 50 μl PCR were electrophoresed through agarose prior to blotting. Blots were probed with the central 1.6-kb HindIII fragment of Ac or a fragment of the FCA gene (Macknight and Dean, unpublished data).

Cloning and sequencing of Ac RT-PCR products

Reverse transcription was carried out as described above. Amplifications were carried out using primers INT1-F and INT4-R
(defined above). Oligo (dT) was also present in the reaction mixture, having been carried over from the cDNA synthesis step. The following reaction conditions were used: 30 sec at 94°C, 30 sec at 55°C and 30 sec at 72°C for 50 cycles; the shortened incubation periods and increased number of cycles were intended to facilitate the selective amplification of smaller products. Amplification was preceded by denaturation at 94°C for 5 min. Products were cloned using the vector pCR2.1 (Invitrogen). Positive clones were identified by colony blotting and hybridization, according to the Hybond N protocol (Amersham); a probe comprising the entire Ac element was utilized. Clones were then sequenced using M13 forward and reverse primers (Li-Cor) and an automated DNA sequencer (Li-Cor).

Acknowledgements

We are extremely grateful to Clare Lister for the screening and sequence analysis of the aberrantly processed Ac transcripts. We thank Steve Scofield and Jonathan Jones for providing the genomic and cDNA Ac clones with the introduced Cael and EcoRV restriction sites, Craig Simpson and John Brown of the Scottish Crop Research Institute for their help and advice regarding the use of quantitative RT–PCR, Tania Page for generating the Arabidopsis transformants and David Young and Billy Sheppard for their assistance in the cloning of the Ac RT–PCR products. The authors acknowledge funding from the Salter Studentship to P.J., Laval University to F.B, and the BBSRC (Grant No. PMB208/0609) to C.D. This research was conducted in accordance with MAFF license PHF 1418A/95234.

References


