Intron retention in Arabidopsis mRNA transcripts.

Craig G Simpson[1], Maria Kalyna[2], John Fuller[1], Diane Davidson[1], Andrea Barta[2] and John WS Brown[1,3]

1. Genetics Programme, SCRI, Invergowrie, Dundee, UK. DD2 5DA.
2. Max F Perutz Laboratories, Medical University of Vienna, Dr. Bohr-Gasse 9/3, A-1030 Vienna, Austria
3. Division of Plant Sciences, University of Dundee @ SCRI, UK DD2 5DA

Introduction

Bioinformatic analysis of genomic and cDNA/EST data have allowed the identification of alternative splicing variants in eukaryotic genes. Alternative splicing data sets for species with relatively low EST coverage and depth of sequence, like Arabidopsis, are incomplete. ESTs can be biased towards the 3'end of transcripts, contain DNA derived artefacts from unspliced or partially spliced mRNAs, or be incomplete because of low levels of expression or rare alternative splicing events in specific cells or developmental stages. Intron retention events, in particular, are prone to genomic DNA contamination or the cloning and sequencing of partially spliced mRNAs.

Intron retention is the most commonly described form of alternative splicing in plants, where introns are small and have a high UA sequence content. Intron and exon definition is known to occur in plants but the size of plant introns, which are significantly smaller than human introns, suggest that intron definition is likely to be more prevalent and errors in splicing may, therefore, be expected to lead to inclusion of the intron in the final mRNA. Plant introns are UA-rich, which increases the likelihood that stop codons are introduced into mRNA transcripts that retain introns and also increases the chances that these transcripts may be targeted for nonsense mediated decay (NMD).

We have developed a high resolution RT-PCR based system that allows us to analyse multiple intron retention events (Simpson et al., 2008). Using this system we were able to compare intron retained transcripts with other alternatively spliced transcripts and evaluate the importance of intron retention in the transcriptome. Using mutants impaired in NMD we further examined whether transcripts containing retained introns were substrates for NMD.

Differences in plant and human intron and gene structure

Average

	Gene size	Exon size	Intron size
Human	28kbp	171bp	5.5kbp
Arabidopsis	2kbp	217bp	167bp

Main AS types

	Exon skip	Alt 5'ss	Alt 3'ss	Intron retention
Human	42%	24%	25%	9%
Arabidopsis	3%	18%	38%	41%

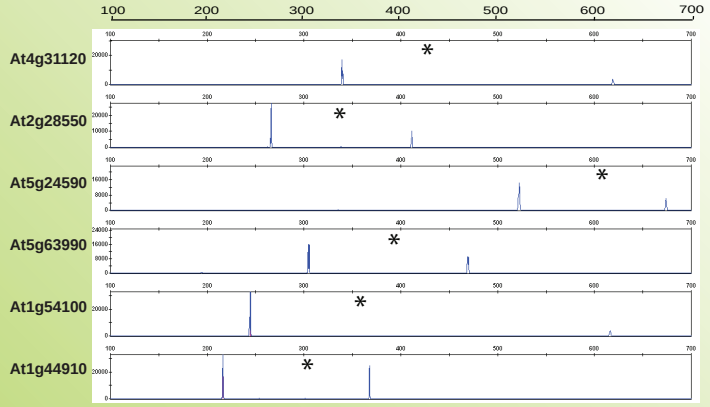
Barbazuk et al., 2008

Summary of intron retention analysis

We used the high resultion RT-PCR alternative splicing panel to anlayse multiple intron retention events (annotated in databases) and determine the significance of the intron retained transcripts to other spliced transcripts. We used 10 day old seedling plants to identify the ratios between transcripts that retain the intron and spliced or alternatively spliced transcripts. We further used seedling plants of mutants that are defecrtive in NMD (*upf-1-5* and *upf 3-1*).

- 46 IR analysed. Primers were designed to cover 2-4 introns.
- 10 expected IR products were not seen at all.
- 11 expected IR products were barely detectable - <1% of the total amount of spliced transcripts.
- 25 expected IR products were detectable between 2 and 100% of the total transcripts of which:
  - 2 were in frame
  - 2 showed no fully spliced product was visible (no intron at all, rarely spliced)
  - 5 changed the C-terminus of protein (NMD not expected)
  - 3 were in 5'UTR. IR leaves uORFs (one showed NMD).
  - 13 are PTC+ but no evidence of NMD to date.

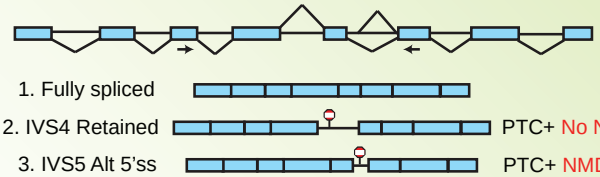
Many intron retention events are not detected



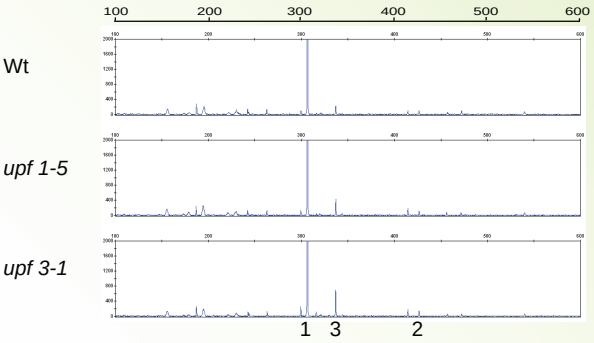
* indicates the position and size of the expected IR transcript. The remaining peaks show spliced and unspliced products. Approximately half of tthe intron retained alternative splicing events analysed gave no or barely detectable transcripts and may, therefore, represent partially spliced mRNAs or very rare ESTs. This suggests that the amount of recognised intron retention in Arabidopsis is likely to be over-estimated.

PTC in IR transcripts are not sensitive to NMD

At5g24270 Calcineurin B-like protein (SOS3)

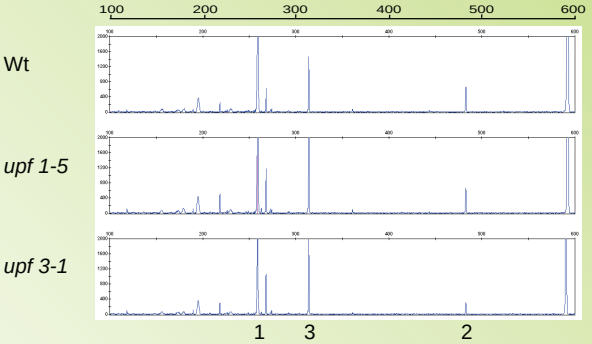


1. Fully spliced
2. IVS4 Retained — PTC+ No NMD
3. IVS5 Alt 5'ss — PTC+ NMD



At1g76460 RNA recognition motif (RRM)-containing protein

1. Fully spliced
2. IVS4 Retained — PTC+ No NMD
3. IVS5 Alt 3'ss — PTC+ NMD



About half of the detectable IR transcripts have a PTC that may lead to a truncated protein. Comparison of spliced and IR transcripts in wild type and mutants of the NMD pathway show no significant change in the ratio of spliced to IR transcript in the NMD mutants. This suggests that intron retained trascripts are insensitive to NMD. This is highlighted in the two examples above that have other alternative splicing events in their 5' (At5g24270) or 3' (At1g76460) splice sites. These alternative splicing events result in the presence of a PTC that is subject to NMD but the IR transcript remains unaffected.