Plant genes for the spliceosomal protein, PRP8

J.I. Hamilton, C.G. Simpson, G.P. Clark, C. McQuade, J.M. Lyon, J.A. Watters & J.W.S. Brown

The majority of plant genes contain intervening sequences (introns) which interrupt the protein-coding sequence. Following transcription into precursor messenger RNA (pre-mRNA), the introns must be removed to generate the mature mRNA, which is then translated into protein in the cytoplasm. Removal of introns is known as splicing and occurs in a large complex of RNA and proteins called the spliceosome. The spliceosome contains in the region of 100 proteins. Many of these proteins are evolutionarily conserved and are known to be essential for the splicing process in yeast and animal systems. Nevertheless, even in such highly conserved biochemical machinery, there are differences between plants and other eukaryotes in terms of components. One such component, which is known from yeast to be essential for splicing, is the spliceosomal protein, PRP8.

One of the exciting areas to emerge in the splicing field over the last 3-4 years has been the discovery of a class of introns which differ from conventional introns. The former, known as AT-AC introns, are removed by a different spliceosome (Fig. 1). The major spliceosome, which removes conventional introns, contains four main components, termed small nuclear ribonucleoprotein particles or snRNPs: U1, U2, U4/U6 and U5. The minor spliceosome, which removes AT-AC introns, contains U11, U12, $U4_{atac}/U6_{atac}$ and U5snRNPs. As a result, the spliceosomes are called U2-dependent spliceosomes and U12-dependent spliceosomes respectively. While many components differ between the two spliceosomes, it is noteworthy that the U5snRNP is utilised in both, and that PRP8 is a U5snRNP protein.

	5' splice site	Branch point	3' splice site
U2	GU	CURAY	AG
U12	AU GU	UCCUUAAC UCCUUAAC	YYCAG YYCAG

Figure 1 Conserved sequences in U2-dependent and U12-dependent introns.

PRP8 was discovered first in yeast, where at 280 kDa, it is one of the largest proteins in the yeast cell. PRP8 has also been isolated from man, *Caenorhabiditis elegans*, *Arabidopsis* and maize. Maize *PRP8* genes were isolated at SCRI on the basis of conserved sequence. The *Arabidopsis PRP8* gene was isolated by colleagues in Oklahoma State University, USA. These represent the only plant *PRP8* genes cloned to date and we have used them to investigate gene number, expression and potential function of the genes.

Genomic and cDNA *PRP8* clones have been isolated from maize (Fig. 2). The maize gene contains 13 exons and 12 introns, spanning 9.8 kb of genomic sequence. The gene would encode a protein of 2,363 amino acids. The *Arabidopsis* PRP8 gene was isolated from a T-DNA insertion mutant, *sus2-1*, which gave an embryo-lethal phenotype. The isolated gene was able to complement the *sus2-1* mutant, showing that it was functionally equivalent. The coding regions of these two plant genes were over 98% identical at the amino acid level. However, they differed in two significant ways. First, the maize N-terminal region was proline-rich, as found in yeast PRP8, while *Arabidopsis* and *C. elegans* PRP8 proteins did not contain this region (Fig. 2). The proline-rich N-terminal region of yeast is essential for splicing, making the difference between the plant PRP8 proteins intriguing. Sec-

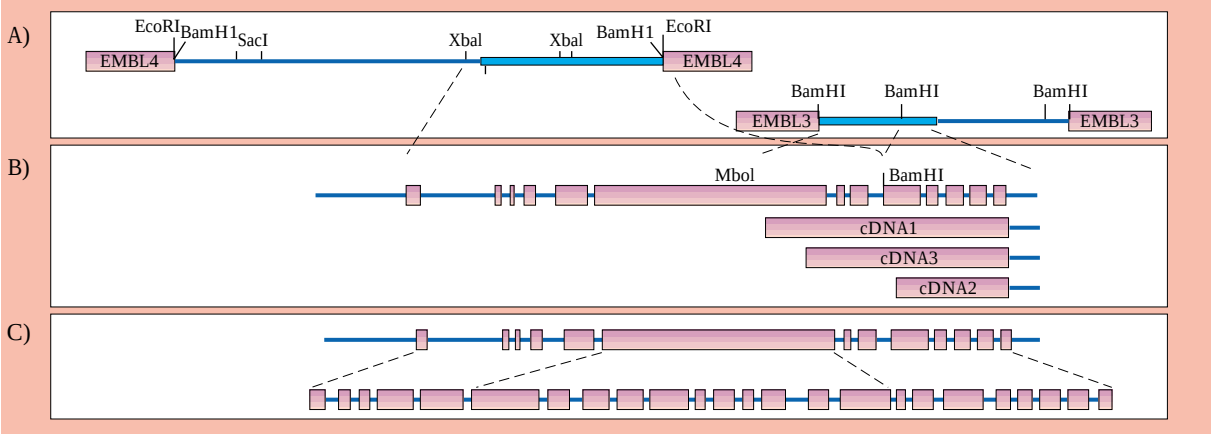

Figure 2 Comparison of gnomic structure of *PRP8* genes isolated from maize and Arabidopsis. Boxes represent exon (coding) sequences which are interspersed by introns (lines)

ond, the *Arabidopsis PRP8* gene contained 23 introns, 11 more than that of maize. Ten of the extra introns in *Arabidopsis* interrupted a region which in maize *PRP8* constituted the large central exon. This finding is contrary to the widely-held view that *Arabidopsis* genes are less complex than those of other plants.

Gene number was analysed by Southern analysis of maize and *Arabidopsis* genomic DNA and by screening *Arabidopsis* yeast artificial chromosome (YAC) libraries. In both, *PRP8* genes are present in small multigene families of 3-5 copies and those in *Arabidopsis* were located on chromosomes 1 and 4. Thus, plants differ from human and yeast systems, which contain only a single *PRP8* gene. That there are 3-5 *PRP8* genes in *Arabidopsis*, but mutation of a single gene in *sus2-1* leads to embryo death, raises further questions on the expression patterns of the different genes or functional diversity. Clearly, in human, the same PRP8 protein will be present in both U2-dependent and U12-dependent spliceosomes. However, the presence of multiple genes in plants, the observed differences between the cloned *Arabidopsis* and maize genes and the phenotype of *sus2-1*, raise the possibility that different PRP8 proteins are found in the U2- and U12-dependent spliceosomes in plants.

At SCRI, we have obtained two lines of evidence to suggest that is the case. First, at least two different *PRP8* genes are expressed in developing embryos at the time when the *sus2-1* embryos abort. Thus, PRP8 protein is likely to be present. Second, in *sus2-1* mutant embryos, splicing of conventional introns is unaffected, suggesting that the *prp8* mutation in *sus2-1* has not affected U2-dependent splicing. Our current hypothesis is that plants contain multiple *PRP8* genes, which encode proteins functioning in either U2- or U12-dependent spliceosomes. The genes are different in exon-intron structure and are assembled into the different spliceosomes on the basis of the presence or absence of the polyproline N-terminal

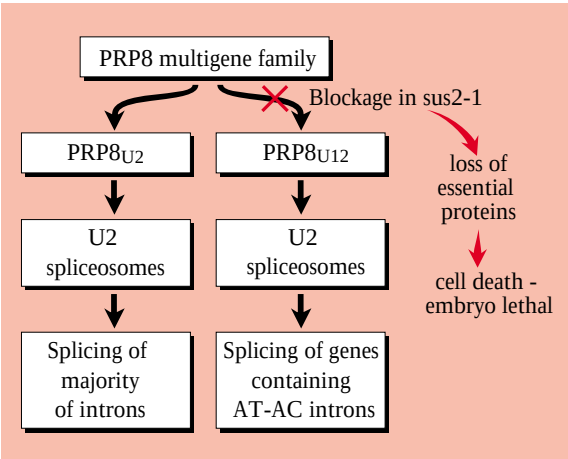

Figure 3 Model for embryo-lethal phenotype of *Arabidopsis sus2-1* mutant.

region. In *sus2-1*, the mutant *prp8* gene may encode a U12-dependent PRP8. If there is only one such gene (reflecting the much lower abundance of U12-dependent spliceosomes), the mutation will knock-out U12-specific PRP8, leading to lack of splicing of all AT-AC intron-containing transcripts. This would lead to embryo death, even in the presence of U2-specific PRP8, U2-dependent spliceosomes, and unaffected splicing of normal introns (Fig. 3). We are currently investigating splicing of AT-AC introns in *sus2-1* mutant embryos and are investigating further the characteristics of individual gene members in the maize and *Arabidopsis* families.