Plant Exon Junction Complex (EJC) proteins

Dominika Lewandowska[1], Sang Hyon Kim[1], Alison Pendle[2], Olga Koroleva[2], Peter J Shaw[2] and John WS Brown[1]

[1] Scottish Crop Research Institute, Invergowrie, Dundee DD2 5DA, Scotland, UK
[2] John Innes Centre, Norwich NR4 7UH, UK.

Scottish Crop Research Institute

Introduction

Splicing of pre-mRNA in higher eukaryotes imprints the resulting mRNA with a multi-protein complex called the exon junction complex (EJC). EJCs are deposited on the transcript in a position-dependant manner: 20-24 nucleotides upstream of the exon-exon junction (Figure 1). Components of the EJC play critical roles in different post-splicing processes like export of mRNA from the nucleus to cytoplasm, the cytoplasmic localisation of mRNA and nonsense-mediated mRNA decay (NMD)/ mRNA surveillance (Dreyfuss et al., 2002; Maquat, 2004). The NMD pathway rapidly degrades mRNAs containing premature translation termination codons and the EJC has been shown to be its key effector in animal cells.

In a recent proteomic analysis of *Arabidopsis* nucleoli we identified six components known from animal studies to be part of the EJC : Aly2, UAP56-2, RNPS1, Y14, Mago and eIF4A-III Pendle et al., 2005). In the *Arabidopsis* genome, we have identified possible orthologues of all EJC proteins except TAP (Table 1).

The nucleolar association was confirmed by GFP-fusion protein localisation for the proteins identified by proteomics and for other EJC components. These results raise the possibility that in plants, nucleoli may have additional functions in mRNA export or surveillance. The protein complement of plant EJCs is unknown, and we are using Tandem Affinity Purification (TAP) to isolate complexes from *Arabidopsis*. Currently, we are using different EJC TAP-tagged proteins expressed in *Arabidopsis* cells as bait to fish for the whole complexes, followed by peptide analysis with mass spectrometry.

Figure 1

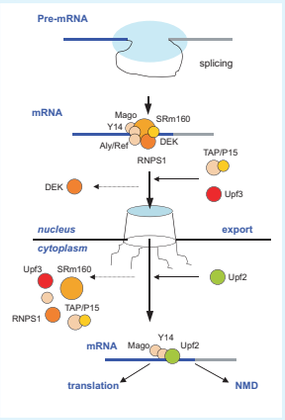

Figure 1. The exon junction complex links transcryption and splicing to mRNA export, translation and NMD (Le Hir et al, 2001; Singh and Lykke-Andersen, 2003)

Table 1

Protein	Number of gene variants	At number
UAP56	2	At5g11200 At5g11170
Upf2	1	At2g39260
Upf3	1	At1g33980
RNPS1	1	At1g16610
SRm160	1	At2g29210
DEK	4	At3g48710 At5g63550 At4g26630 At5g55660
P15	3	At1g11570 At1g27970 At1g27310
Aly/Ref	4	At5g59950 At5g02530 At1g66260 At5g37720
Y14	1	At1g51510
Mago	1	At1g02140
eIF4A-III	1	At3g19760
Pinin	1	At1g15200
MLNS1	1	At1g80000
TAP	0	

Table 1. Exon Junction Complex homologues found in *Arabidopsis thaliana* genome

The nucleolus is the most prominent body in the nucleus (Figure 2). It is multifunctional being involved in a number of aspects of RNA metabolism and RNP assembly besides its classical role in rRNA transcription and processing and ribosomal subunit assembly.

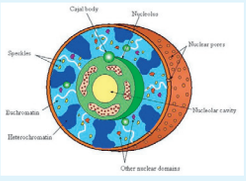

Figure 2. Organisation of the plant nucleus showing nucleolar sub-structure

GFP-Fusion Constructs And Localisation OF 6 EJC Proteins

Full-length cDNAs of the EJC proteins identified by proteomics (Aly2, UAP56-2, RNPS1, Y14, Mago and eIF4A-III) were cloned as N-terminal GFP-protein fusions and expressed from the Cauliflower Mosaic Virus (CaMV) 35S promoter. The GFP-fusions were expressed transiently in *Arabidopsis* cell cultures using *Agrobacterium* infection. GFP fluorescence was successfully detected in all clones and expression patterns were imaged using either confocal microscopy or wide-field CCD imaging followed by image deconvolution. For each fusion, a range of cells with relatively low levels of fluorescence was examined and localisation determined. Weakly fluorescing cells were imaged in order to minimise the possibility of artefacts due to over-expression.
All six proteins show nucleolus-associated labelling as well as a range of different localisation patterns of nucleolar and nucleoplasmic labelling (figure 3).

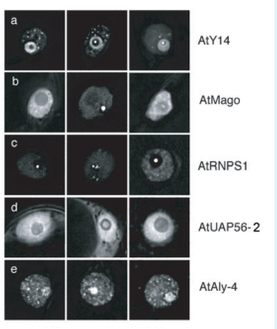

Figure 3. GFP localisations of EJC proteins identified in the Arabidopsis nucleolar proteomic analysis. Images of three different cells are shown for each of the Y14, Mago, RNPS1, UAP56-2 and ALY4 GFP fusions.

Other EJC Components Also Localise To The Nucleolus

To investigate whether this association with the nucleolus was a general phenomenon for Arabidopsis EJC proteins, we prepared GFP fusions for other Arabidopsis homologues of EJC components: the second UAP56 gene (At*UAP56-1)*, the three remaining ALY genes (At*ALY-1*, At*ALY-2* and At*ALY-3*), two P15 genes (AtP15-2 and AtP15-3) one gene of DEK, SRm160 and Upf3. No homologue of TAP was identified.
All fusions except p15-2 showed some labelling of the nucleolus and variable labelling of the nucleoplasm and nuclear bodies (figure 4). In some cases (eg. ALY proteins) there was differential labelling. Thus, in plants all *Arabidopsis* EJC components analysed show some localisation to the nucleolus, and the majority label the nucleolus strongly. This suggests that the nucleolus may be a site of storage, assembly or function of the EJC components or complexes.

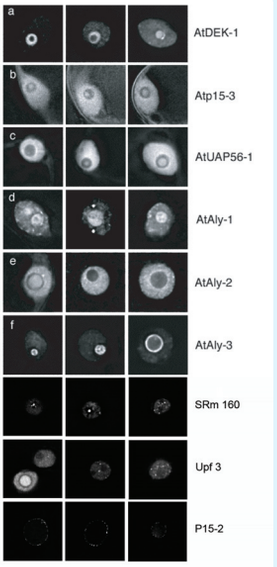

Figure 4. GFP localisations of EJC proteins not detected in the *Arabidopsis* nucleolar proteomic analysis. Images of three different cells are shown for each of the DEK-1, p15-3, UAP56-1, ALY1, ALY2 and ALY3, SRm160, Upf3 and p15-2 GFP fusions.

EJC Proteins Are Also Present In The Human Nucleolar Proteome

In light of the clear association of plant EJC proteins with the nucleolus, we searched the human nucleolar proteome of 692 proteins, and found six EJC components (ALY/Ref, UAP56, Y14, DEK, eIF4A-III and UPF3X) (Andersen et al., 2004). In human and Drosophila cells, several EJC components have been shown to be localised to the nucleoplasm and to nuclear speckles but the current evidence suggests that they are excluded from the nucleolus. Therefore, it may be that the human proteins are only present in the nucleolus in low levels or under particular conditions.

Analysing The Protein Complement Of EJC In Plants

To isolate exon junction complexes from *Arabidopsis* we are using Tandem Affinity Purification (TAP). Full-length cDNAs for all *Arabidopsis* EJC proteins have been cloned into the TAP-tag binary vector as C-terminal protein fusions (Figure 5A) and transiently expressed from the Cauliflower Mosaic Virus (CaMV) 35S promoter in *Arabidopsis* cell cultures. Different EJC TAP-tagged proteins were used as bait to fish for the complexes, followed by peptide analysis with mass spectrometry. So far we have attempted to TAP-purify EJCs with P15-3, Mago, UAP56-1 and eIF4A-III (Figure 5B) from *Arabidopsis* cells. The proteomic analysis by mass spectrometry of purified UAP56-1 and eIF4A-III revealed the presence of bait protein, but to date we were unable to identify any interacting proteins.

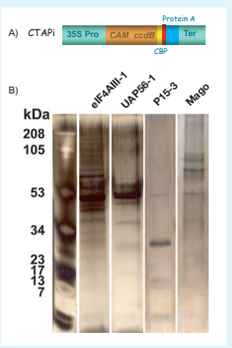

Figure 5. Analysing the protein complement of EJC in plant. A) Binary C-TAPi vector (Rohila et al., 2004) into which cDNAs of all *Arabidopsis* EJC protein were cloned

B) Protein samples after TAP purification from *Arabidopsis* culture cells analyzed by SDS/PAGE (silver stained gels)

Conclusions

1.) Six Exon Junction Complex components were found in the *Arabidopsis* nucleolar proteome and all show association with the nucleolus by GFP-fusions
2.) The *Arabidopsis* genome contains possible orthologues of all EJC protein except TAP
3.)Some of the EJC proteins have multiple gene variants and form small multi-gene families
4.) The four plant ALY proteins and two plant P15 protein variants show differential sub-nuclear localisation suggesting different functions
5.) The plant nucleolus may function in storage or assembly of EJC sub-complexes, mRNA export or in NMD

Acknowledgements and References

This research was supported by the Scottish Executive Environment and Rural Affairs Department and the Biotechnology and Biological Sciences Research Council; DL is an EU Marie Curie Fellow.

Andersen et al., 2005, Nature, 433(7021):77-83
Dreyfuss et al., 2002, Nat Rev Mol Cell Biol. 3(3):195-205
Le Hir et al, 2001, EMBO J. 20(17): 4987-4997
Maquat, 2004, Nat Rev Mol Cell Biol. 5(2): 89-99
Pendle et al., 2005, Mol Biol Cell. 16(1):260-9
Rohila et al., 2004, Plant J. 38(1): 172-81
Singh and Lykke-Andersen, 2003, TIBS 28(9): 464-466