Nuclearly Encoded Splicing Factors Implicated in RNA Splicing in Higher Plant Organelles

Andéol Falcon de Longevialle a,b, Ian D. Small b and Claire Lurina a

a Unité Mixte de Recherche en Génomique Végétale (Institut National de la Recherche Agronomique/Centre National de la Recherche Scientifique/Université d’Evry Val d’Essonne), 91057 Evry, France
b Australian Research Council of Excellence in Plant Energy Biology, University of Western Australia, Crawley 6009, WA, Australia

ABSTRACT Plant organelles arose from two independent endosymbiosis events. Throughout evolutionary history, tight control of chloroplasts and mitochondria has been gained by the nucleus, which regulates most steps of organelle genome expression and metabolism. In particular, RNA maturation, including RNA splicing, is highly dependent on nuclearly encoded splicing factors. Most introns in organelles are group II introns, whose catalytic mechanism closely resembles that of the nuclear spliceosome. Plant group II introns have lost the ability to self-splice in vivo and require nuclearly encoded proteins as cofactors. Since the first splicing factor was identified in chloroplasts more than 10 years ago, many other proteins have been shown to be involved in splicing of one or more introns in chloroplasts or mitochondria. These new proteins belong to a variety of different families of RNA binding proteins and provide new insights into ribonucleoprotein complexes and RNA splicing machineries in organelles. In this review, we describe how splicing factors, encoded by the nucleus and targeted to the organelles, take part in post-transcriptional steps in higher plant organelle gene expression. We go on to discuss the potential for these factors to regulate organelle gene expression.

Key words: RNA splicing; chloroplast; mitochondria; splicing factors.

INTRODUCTION

Eukaryotic cells contain organelles that are essential for supplying energy via ATP synthesis, namely mitochondria and in photosynthetic cells, chloroplasts. These two organelles are the results of two endosymbiosis events: mitochondria derive from endosymbiotic α-proteobacteria (Andersson et al., 2003) and chloroplasts derive from endosymbiotic cyanobacteria (Raven and Allen, 2003). Throughout evolution, organelle genomes have lost much of their original genomes by the redistribution of genetic material between nucleus, mitochondria, and plastids via inter-compartmental DNA transfer (Ayliffe et al., 1998; Ellis, 1982; Stern and Lonsdale, 1982; Weeden, 1981). From the six theoretical DNA transfer possibilities, at least five have been observed (Kleine et al., 2009): DNA transfer from chloroplasts to mitochondria and, vice versa, DNA transfer from nucleus to mitochondria and the two most common events, DNA transfers from chloroplasts or mitochondria to the nucleus. Many of the proteins encoded by genes transferred from organelles to the nucleus are important for organelle gene expression or metabolism and need to be targeted back to their original compartment, creating a requirement for coordinate regulation of nuclear and organelle gene expression. In addition to the genes transferred from organelles, many other nuclearly encoded proteins have acquired functions in different processes of gene expression in organelles. These host-encoded proteins are needed for a wide range of transcriptional and post-transcriptional processes including DNA replication, RNA transcription, RNA processing, RNA editing, RNA splicing and translation. RNA splicing is a good illustration of the typical evolutionary trajectory of these processes, which all illustrate a steadily increasing role for the cell nucleus in controlling organelar activities. Originally, RNA splicing of introns was enhanced by their intron-encoded proteins (IEPs). These contained endonuclease and reverse transcriptase domains believed to play an enzymatic role in genetic mobility and a maturase domain, thought to promote splicing. Introns were dispersed as mobile elements, but in land plants, almost all introns have lost their IEPs, ending with only two maturases encoded in the organelar genomes and a set of nuclearly encoded splicing factors.

1 To whom correspondence should be addressed. E-mail lurin@evry.inra.fr, fax 33 (0)1 60 87 45 10, tel. 33 (0)1 60 87 45 13.
© The Author 2010. Published by the Molecular Plant Shanghai Editorial Office in association with Oxford University Press on behalf of CSPP and IPPE, SIBS, CAS.
Received 12 March 2010; accepted 27 April 2010
INTRON STRUCTURE AND SPlicing MECHANISMS

RNA splicing is a processing event in which the introns of a precursor messenger RNA (pre-mRNA) are removed and its exons are joined. In plants, several chloroplast and mitochondrial genes, encoding either tRNAs or proteins, are interrupted by introns. Depending on their conserved primary and secondary structures as well as their different splicing mechanisms, these introns are classified into two main families, group I and group II introns. Twenty group II introns and only one group I intron have been found in the chloroplast genome of *Arabidopsis thaliana* while 17 group II and one group I intron have been described in maize and rice. Similarly, plant mitochondrial genomes contain mostly group II introns. Among flowering plant mitochondrial genomes, 25 group II introns have been described (Bonen, 2008), most of which disrupt nad genes. In addition, one group I intron, coding an intronic ORF with potential maturase function, was found in the cox1 gene of several plants (Cho et al., 1998).

Group I introns are characterized by a secondary structure of 10 domains, P1 to P10, each of them having a specific role in RNA folding to enhance splicing and ligation. Splicing of group I intron is mediated via a two-step phosphoryl transfer reaction using an exogenous guanosine as a cofactor instead of the adenosine in domain VI used in group II introns (Kruger et al., 1982).

Group II introns are composed of six domains (Figure 1) and are divided into four subclasses. Two subclasses, group IIA and IIB, are found in flowering plants (Michel et al., 1989), while two additional subclasses, group IIC and IID, have been discovered only in bacteria (Michel et al., 1989; Toor et al., 2001). For the most part, the overall organization and structure are similar in all these classes and reviews of their structure and folding have been published (Fedorova et al., 2007; Keating et al., 2010; Pyle et al., 2007). Domain I, which is essential for ribozyme activity, contains two exon binding sites (EBS1 and EBS2) that interact with two corresponding intron binding sites (IBS1 and IBS2) located at the 3' end of the first exon. Domains II and III enhance the catalysis of splicing, while domain IV, which contains the coding sequence for a maturase in the case of nad1 intron 4 in plant mitochondria and trnK intron in chloroplasts, is not involved in catalysis. This domain is also the break point for eight events of trans-splicing that have been observed in plants (rps12 intron 1 in chloroplasts, nad1 intron 1, intron 3, intron 4; nad2 intron 2; nad5 intron 2, intron 3 in mitochondria (Bonen, 2008), and recently in the mitochondrial cox2 intron in *Allium cepa* L. (Kim and Yoon, 2010)). Domain V is the most phylogenetically conserved domain and contains a sequence of 34 nucleotides that forms part of the catalytic core and is essential for ribozyme activity. Domain VI provides the branchpoint adenosine for the first trans-esterification step. Although these features are common to most group II introns in plants, some exceptions have been described. For Figure 1. Scheme of the Secondary Structure of Group II Introns and their Splicing Mechanisms. The catalytic mechanism of group II introns closely resembles that of the spliceosome. Splicing occurs via two trans-esterification steps, annotated 1 and 2. First, the 2' OH of the branch point adenosine within domain VI performs a nucleophilic attack on the first nucleotide of the intron at the 5' splice site, forming the lariat intermediate. Second, the 3' OH of the released 5'-exon performs a nucleophilic attack at the last nucleotide of the intron at the 3' splice site, joining the exons and forming a lariat that will be polyadenylated and degraded.
example, nad1 intron 1, which does not have a bulged adenosine but, in contrast, has an aberrantly large domain V loop (Carrillo et al., 2001; de Longevialle et al., 2007); nad1 intron 2, which has a very short, strongly base-paired domain VI helix (Li-Pook-Than and Bonen, 2006); nad4 intron 2, which lacks a bulged adenosine at the expected position (Bonen, 2008); ycf3 intron 2, which has two additional nucleotides in addition to the 34 nucleotides expected in domain V (de Longevialle et al., 2008); and the chloroplast trnV intron, which lacks a bulged A in domain VI and is spliced by a hydrolytic mechanism (Vogel and Borner, 2002).

In most species, many of these group I or group II introns are mobile genetic elements whose mobility is mediated by a ribonucleoprotein (RNP) complex composed of the intron RNA and a conserved intron-encoded protein, also called maturase, which facilitates the splicing of its host intron (Eickbush, 1999; Wank et al., 1999). In flowering plants, most of the organellar intron maturases have been lost. The only remaining maturases are found in the trnK intron of the chloroplasts, in nad1 intron 4 of mitochondria and in the cox1 intron of several plants (Cho et al., 1998). The loss of these maturases has been accompanied by the recruitment of host-encoded proteins to facilitate splicing of introns that have lost their autocatalytic activity. None of the group II introns in plants that have been tested was able to self-splice in vivo and all have been shown to require additional trans factors for efficient splicing in vivo.

The two first nuclear mutations affecting a subset of chloroplast splicing events were identified in maize in the crs1 and crs2 mutants (described below), opening a new research area to elucidate the mechanisms that regulate organellar RNA splicing (Jenkins et al., 1997).

In mitochondria, the first nuclear mutation affecting a splicing event was identified in Nicotiana sylvestris (Brangeon et al., 2000) in the nuclear nms1 mutant. This mutant presented a defect in RNA splicing of nad4 intron 1 and no spliced nad4 exon1-exon2 could be detected while all other transcripts were not affected. Unfortunately, the Ms1 locus has never been identified and thus no information on the Ms1 protein is available.

Since then, many other factors acting as splicing factors to facilitate or enhance RNA splicing have been molecularly characterized during the last 10 years. These factors belong to different families of RNA binding proteins that have acquired specific roles in RNA splicing during the evolution of land plants. Their roles are summarized in Table 1, Figure 2, and Figure 3 and each of them will be discussed in detail below.

Maturases

Maturase proteins are divided into two classes based on the location of their genes, intron-encoded maturases, and nucleolar encoded maturases. Each organelle genome in plants has conserved only a single group II intron-encoded maturase. In chloroplasts, MatK is encoded within the intron of trnK and in mitochondria, MatR is encoded within the fourth intron of nad1. The position of these maturase genes inside introns suggests a specific role in the splicing of precursor RNAs transcribed from the genes in which they reside, but until very recently, no evidence for this was available. Immunoprecipitation of MatK-associated RNA (Zoschke et al., 2010) captured seven out the eight group IIA introns in tobacco chloroplasts, including the trnK intron. MatK preferentially associated with domains II and IV of the trnK intron. These co-purifications strongly suggest a role of MatK in RNA splicing for these seven introns. A similar approach would be very helpful to elucidate the role of MatR in mitochondria. The second class of maturase was discovered by analyzing the complete nuclear genomes of Arabidopsis and rice. Four nuclear genes coding for proteins closely related to group II intron-encoded maturases were found in both genomes (Mohr and Lambowitz, 2003). Three maturases are targeted to mitochondria (Table 1) and one (At1g74350) has a dual localization, depending on the start codon used for its translation (Keren et al., 2009). The longer isoform was observed in mitochondria, while the shorter one is targeted to chloroplasts. All these maturases are presumed to act in trans in the splicing of organellar introns. To date, a role in RNA splicing has been demonstrated by reverse genetic approaches for only two out of the four nuclearly encoded maturases. At-nMat1a (At1g30010) is involved in the matura-

CRM Domain Proteins

The Chloroplast RNA splicing and ribosome maturation (CRM) domain was first described in RNA splicing factor CRS1 (Jenkins et al., 1997; Till et al., 2001). Although this domain is found in at least five characterized chloroplast splicing factors in plants, it derived from an ancient ribosome-associated protein (Barkan et al., 2007). CRM domains in prokaryotes exist as stand-alone proteins of ∼100 amino acids encoded by a single copy ORF. The Escherichia coli YhbY protein has been shown to be associated with pre-50S ribosomal subunits, suggesting a function in ribosome assembly. In contrast to prokaryotes, plant genomes encode multiple CRM domain proteins. Sixteen Arabidopsis and 14 rice proteins containing one or more CRM domains...
<table>
<thead>
<tr>
<th>Type</th>
<th>Name</th>
<th>A. thaliana</th>
<th>O. Sativa</th>
<th>Sub-Type</th>
<th>Localisation</th>
<th>Affected introns</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>maturase proteins</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AtnMat3</td>
<td>At5g04050</td>
<td>OS06G44880</td>
<td>maturase</td>
<td>mitochondria</td>
<td>unknown</td>
<td>Keren et al., 2009</td>
</tr>
<tr>
<td></td>
<td>AtnMat4</td>
<td>At1g74350</td>
<td>OS06G42770</td>
<td>maturase</td>
<td>mitochondria</td>
<td>unknown</td>
<td>Keren et al., 2009</td>
</tr>
<tr>
<td></td>
<td>At-nMat1a</td>
<td>At1g30010</td>
<td>OS12G21870</td>
<td>maturase</td>
<td>mitochondria</td>
<td>nad4</td>
<td>Nakagawa and Sakurai, 2006</td>
</tr>
<tr>
<td></td>
<td>AtnMat2</td>
<td>At5g46920</td>
<td>OS10G41790</td>
<td>maturase</td>
<td>mitochondria</td>
<td>cox2 intron, nad1 intron2, nad7 intron2</td>
<td>Keren et al., 2009</td>
</tr>
<tr>
<td></td>
<td>matK</td>
<td></td>
<td></td>
<td>maturase</td>
<td>chloroplasts</td>
<td>atpF intron, trnK intron, trnA intron, trnF intron, trnV intron, rpl2 intron, rps12 intron</td>
<td>Zoschke et al., 2010</td>
</tr>
<tr>
<td><strong>matR</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRM proteins</td>
<td>AtCFM2</td>
<td>At3g01370</td>
<td>Os04390800</td>
<td>CRS1 subfamily</td>
<td>chloroplasts</td>
<td>trnL intron, ndhA intron, ycf3 intron1, clpP intron2</td>
<td>Asakura and Barkan 2007</td>
</tr>
<tr>
<td></td>
<td>AtCFM3a</td>
<td>At3g23070</td>
<td>Os11g37990</td>
<td>chloroplasts and mitochondria</td>
<td>unknown</td>
<td></td>
<td>Asakura et al., 2008</td>
</tr>
<tr>
<td></td>
<td>AtCFM3b</td>
<td>At4g14510</td>
<td></td>
<td>chloroplasts</td>
<td></td>
<td></td>
<td>Asakura et al., 2008</td>
</tr>
<tr>
<td></td>
<td>ZmCRS1</td>
<td>At5g16180</td>
<td>Os08g27150</td>
<td>chloroplasts</td>
<td></td>
<td>atpF intron</td>
<td>Jenkins et al., 1997; Till et al., 2001; Asakura and Barkan 2006</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>At3g18390</td>
<td>Os05g47850</td>
<td></td>
<td>chloroplasts</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>At4g29750</td>
<td>Os09g19850</td>
<td></td>
<td>chloroplasts</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AtCFM6</td>
<td>At4g13070</td>
<td>Os06g20030</td>
<td>Subfamily 3</td>
<td>mitochondria and nucleus</td>
<td>unknown</td>
<td>Barkan et al., 2007</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>At2g28480</td>
<td>Os01g72820</td>
<td></td>
<td>chloroplasts</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>At3g25440</td>
<td>Os04g41550</td>
<td></td>
<td>chloroplasts</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>At3g27550</td>
<td>Os05g05300</td>
<td></td>
<td>chloroplasts</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>At5g54890</td>
<td>Os08g08860</td>
<td></td>
<td>chloroplasts</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>At4g31010</td>
<td>Os08g07790</td>
<td></td>
<td>chloroplasts</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZmCAF2</td>
<td>At1g23400</td>
<td>Os01g21990</td>
<td></td>
<td>chloroplasts</td>
<td></td>
<td>ndh8 intron, petB intron, rps12 intron1, ndhA intron and ycf3 intron1*</td>
<td>Ostheimer et al., 2003; Asakura and Barkan 2006</td>
</tr>
<tr>
<td>ZmCAF1</td>
<td>At2g20020</td>
<td>Os01g31110</td>
<td></td>
<td>chloroplasts</td>
<td></td>
<td>petD intron, rpl16 intron, rps16 intron, trnG intron, ndhA intron, ycf3 intron1, rpoC1 intron* and ClpP intron1*</td>
<td>Ostheimer et al., 2003; Asakura and Barkan 2006</td>
</tr>
<tr>
<td></td>
<td>At2g21350</td>
<td>Os10g36860</td>
<td></td>
<td>chloroplasts or mitochondria</td>
<td>unknown</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>At4g39040</td>
<td></td>
<td></td>
<td>chloroplasts</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table 1. RNA Splicing in Plant Organelles

<table>
<thead>
<tr>
<th>Type</th>
<th>Name</th>
<th>A. thaliana</th>
<th>O. Sativa</th>
<th>Sub-Type</th>
<th>Localisation</th>
<th>Affected introns</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PPR proteins</strong></td>
<td><strong>ZmPPR4</strong></td>
<td>At5g04810</td>
<td>Os04g58780</td>
<td>Pure subfamily</td>
<td>chloroplasts</td>
<td>trans rps12 intron1</td>
<td>Schmitz-Linneweber et al., 2006</td>
</tr>
<tr>
<td></td>
<td><strong>AtOTP51</strong></td>
<td>At2g15820</td>
<td>Os02g47360</td>
<td>pure subfamily</td>
<td>chloroplasts</td>
<td>ycf3 intron2</td>
<td>Falcon de Longevialle et al., 2008</td>
</tr>
<tr>
<td></td>
<td><strong>ZmPPR5</strong></td>
<td>At4g39620</td>
<td>Os02g51480</td>
<td>pure subfamily</td>
<td>chloroplasts</td>
<td>trnG-UCC intron</td>
<td>Beick et al., 2008; Cushing D.A. et al., 2005</td>
</tr>
<tr>
<td></td>
<td><strong>AtHcf152</strong></td>
<td>At13g09650</td>
<td>Os12g0102600</td>
<td>or Os11g0103000</td>
<td>chloroplasts</td>
<td>pet8 intron</td>
<td>Meierhoff et al., 2003</td>
</tr>
<tr>
<td></td>
<td><strong>AtOTP43</strong></td>
<td>At1g74900</td>
<td>Os02g45590</td>
<td>pure subfamily</td>
<td>mitochondria</td>
<td>trans nad1 intron1</td>
<td>Falcon de Longevialle et al., 2007</td>
</tr>
<tr>
<td><strong>peptidyl-tRNA hydrolase</strong></td>
<td><strong>ZmCRS2</strong></td>
<td>At5g38290</td>
<td>?</td>
<td>chloroplasts</td>
<td>rps12 intron1, ndhB intron, pet8 intron, ndhA intron, ycf3 intron1, rpl16 intron, rps16 intron, petD intron, trnG intron, clpP intron1* and rpoC1 intron*</td>
<td>Jenkins et al., 2001; Jenkins et al., 1997</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>At5g16140</strong></td>
<td>At5g16140</td>
<td>Os01g49900</td>
<td>os05g47630</td>
<td>not predictable</td>
<td>unknown</td>
<td>unknown</td>
</tr>
<tr>
<td></td>
<td><strong>At1g18440</strong></td>
<td>At1g18440</td>
<td>Os01g49900</td>
<td>or Os05g47630</td>
<td>mitochondria</td>
<td>unknown</td>
<td>unknown</td>
</tr>
<tr>
<td></td>
<td><strong>At5g19830</strong></td>
<td>At5g19830</td>
<td>?</td>
<td>mitochondria</td>
<td>unknown</td>
<td>unknown</td>
<td>unknown</td>
</tr>
<tr>
<td><strong>ribonuclease III proteins</strong></td>
<td><strong>ZmRNC1</strong></td>
<td>At4g37510</td>
<td>Os01g59510</td>
<td>chloroplasts</td>
<td>trnV intron, trnK intron, trnL intron, trnA intron, atpF intron, rps12 intron2, ndhB intron, pet8 intron, petD intron, trnG intron</td>
<td>Watkins et al., 2007</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>DUF860 or PORR</strong></td>
<td>ZmWTF1</td>
<td>At4g01037</td>
<td>Os05g49610</td>
<td>chloroplasts</td>
<td>trnV intron, trnK intron, trnL intron, trnA intron, atpF intron, rps12 intron2, ndhB intron, pet8 intron, petD intron, trnG intron</td>
<td>Kroeger et al., 2009</td>
</tr>
<tr>
<td></td>
<td><strong>Whirly</strong> protein</td>
<td>ZmWHY</td>
<td>At1g14410 or At2g02740</td>
<td>Os06g05350</td>
<td>chloroplasts</td>
<td>atpF intron</td>
<td>Prikril et al., 2008</td>
</tr>
<tr>
<td></td>
<td><strong>DEAD-box protein</strong></td>
<td>AtPMH2</td>
<td>At13g22330</td>
<td>OS12G41715</td>
<td>mitochondria</td>
<td>nad1 intron 2 and 3, nad2 intron 1, 2 and 4, nad4 intron2 and 3, nad5 intron 1, 2 and 3, nad7 intron1 and 4, cox2 intron, rps3 intron and rpl2 intron</td>
<td>Matthes et al., 2007; Kohler et al., 2009</td>
</tr>
<tr>
<td></td>
<td><strong>unknown</strong></td>
<td>NsNMS1</td>
<td>unknown</td>
<td>unknown</td>
<td>mitochondria</td>
<td>nad4 intron1</td>
<td>Brangeon et al., 2000</td>
</tr>
</tbody>
</table>

Yellow boxes represent the species in which the different splicing factors have been characterized. Intron found in *Arabidopsis thaliana* but not in maize and rice are marked with asterisks. Orthologous genes in *Oryza sativa* or in *Arabidopsis thaliana* have been determined using FLAGdb++ v4.2 (http://urgv.evry.inra.fr/FLAGdb) or BLASTP (http://blast.ncbi.nlm.nih.gov/Blast.cgi). The intron whose splicing requires AtCFM3a is marked in red. No clear orthologs of ZmCRS2 and At5g19830 were found in *Oryza sativa*. Sub-cellular localizations confirmed by reverse genetics analysis and/or by localization studies using a fluorescent protein fusion are indicated in bold.
domains have been described and were placed into 14 orthologous groups divided into four subfamilies (Barkan et al., 2007). Seven CRM domain proteins are predicted to be targeted to chloroplasts and among them, five have been characterized as plastidial splicing factors: CFM3, CFM2, CRS1, CAF1, and CAF2 (described below). Two CRM proteins are predicted to be mitochondrial and five others are potentially targeted to two locations: mitochondria and chloroplasts, or mitochondria and nucleus (Table 1).

All the CRM proteins characterized so far are involved in RNA splicing in chloroplasts. The first CRM gene, CRS1 (for Chloroplast RNA Splicing 1), was identified from a genetic screen in maize (Jenkins et al., 1997) and later cloned (Till et al., 2001). CRS1 contains three CRM domains and is required for the splicing of the group IIA atpF intron via a direct interaction with its intronic target. Two other CRM domain proteins were discovered in a yeast two-hybrid screen using CRS2 (a homolog of peptidyl-tRNA hydrolases, which probably lacks that catalytic activity but is required for the splicing of a subset of group II introns) as a bait (Ostheimer et al., 2003). These two proteins, CAF1 and CAF2 (for CRS2-associated factors 1 and 2), have two CRM domains and each of them is bound to CRS2 in chloroplasts. The complex CRS2–CAF1 is required for the splicing of a subset of subgroup IIB introns in maize: petD intron, rpl16 intron, rps16 intron, trnG intron, ycf3 intron 1 (Ostheimer et al., 2003). In Arabidopsis, two additional introns, absent in maize (poc1 intron and ClpP intron 1), are spliced by the CRS2–CAF1 complex (Asakura and Barkan, 2006). The complex CRS2–CAF2, whose formation is mediated by a 22-amino-acid motif in the C-terminal region of CAF2 (Ostheimer et al., 2006), is required for the splicing of five subgroup IIB introns: ndhB intron, petB intron, rps12 intron 1, ndhA intron, and ycf3 intron 1. It was observed that for this last intron, CRS2–CAF1 and CRS2–CAF2 complexes have overlapping functions in maize and in Arabidopsis thaliana (Asakura and Barkan, 2006; Ostheimer et al., 2003). Two Arabidopsis paralogs (At4g31010 and At5g54890) of CAF1 and CAF2 were predicted to be mitochondrial (Table 1). These proteins end just after the second CRM domain, lacking the C-terminal CRS2-binding peptide found in CAF1 and CAF2. In addition, At5g19830, a paralog of CRS2, is also predicted to be mitochondrial and does not contain the conserved residues shown in CRS2 to bind.

Figure 2. Schematic Representation of Chloroplast Splicing Factors in Land Plants. The group II introns are divided into two subgroups IIA and IIB according to Michel et al. (1989). The introns that are found in Arabidopsis thaliana but not in maize and rice are marked with asterisks. PPR proteins are represented in gray, CRM domains proteins in green, peptidyl-tRNA hydrolase protein in blue, maturase in gold, ribonuclease III domains protein in brown, PORR domain protein in red, and ‘Whirly’ protein in purple. This figure is modeled after the figure presented in Kroeger et al. (2009).
CAF1 and CAF2. Nevertheless, it has another hydrophobic domain that could potentially bind to CAF paralogs (Ostheimer et al., 2005). It seems likely that CRS2–CAF-type complexes could exist in mitochondria, where they may be involved in RNA splicing (Ostheimer et al., 2006).

Another CRM domain protein found to be involved in RNA splicing is closely related to CRS1 but has a fourth CRM domain at the C-terminus. CFM2 (for CRM Family Member 2) was found in the large RNP complexes that contain CAF1 and/or CAF2. RNA co-immunoprecipitation assays (RIP-Chip) in maize have shown that CFM2 is associated with one group I intron (trnL intron) and two group II introns (ndhB intron, ycf3 intron 1). In Arabidopsis, a reverse genetic approach has shown that CFM2, in addition to these three introns, also promotes the splicing of the clpP intron 2 (a group II intron), which is not found in maize (Asakura and Barkan, 2007; Asakura et al., 2008).

CFM3 (for CRM Family Member 3), the latest CRM domain protein to be characterized in plants, contains three CRM domains (Asakura et al., 2008). In contrast to the other splicing factors in the family, CFM3 is dual-targeted to chloroplasts and mitochondria. In chloroplasts, CFM3 associates in vivo with a distinct subset of CAF/CRS2-dependent group II introns (ndhB intron, rpl16 intron, rps16 intron, trnG intron, petB intron, and petD intron) and promotes their splicing in rice. In Arabidopsis thaliana, two paralogs have been described: AtCFM3a (At3g23070) is dual-targeted to both organelles and AtCFM3b (At4g14510) is specific to chloroplasts. Loss of AtCFM3a leads to a strong decrease in the splicing of ndhB intron while atcfm3b mutants do not have any phenotype. The double mutant atcfm3a/atcfm3b presents seed lethality, demonstrating that these two proteins are functionally redundant (Asakura et al., 2008) and may indirectly be involved in translation by having a function in splicing of rpl16 (encoding a ribosomal protein) or the trnG intron, as has been demonstrated for OsCFM3 (Asakura et al., 2008). In mitochondria, the role of CFM3 is not completely elucidated and further study will be required.

Among the CRM domain family, four additional proteins cluster in the subfamily 3 (Table 1) (Barkan et al., 2007). These CRM domain proteins have not been characterized so far but are predicted to localize to the nucleus and mitochondria, and this has been confirmed by GFP fusion in the case of CFM6 (At4g13070) (Barkan et al., 2007). It has been suggested that CFM6 may function in ribosome biogenesis within the nucleus and mitochondrion. However, an additional function in RNA splicing in mitochondria is not excluded, as also suggested for the mitochondrial paralogs of CAF1 and CAF2.

Thus, CRM domain proteins are implicated in the metabolism of three classes of catalytic RNA: group I introns, group II introns, and 23S rRNA.

PPR Proteins
Pentatricopeptide repeat (PPR) proteins are RNA-binding proteins that have been identified from the Arabidopsis thaliana genome sequence and are defined by a canonical 35-amino-acid motif, repeated in tandem up to 30 times (Small and Peeters, 2000). More than 450 members compose this family in each land plant genome while only a few per genome have been described in mammals, suggesting multiple specific roles in plants (Schmitz-Linneweber and Small, 2008). Plant PPR proteins can be classified into two major subfamilies based on the nature of their PPR motifs: the P (for pentatricopeptide) subfamily and the PLS subfamily (where L and S indicate long and short variants of the PPR motif). The latter subfamily can be further divided into several subclasses based on their characteristic C-terminal motifs (Lurin et al., 2004). Abundant genetic evidence suggests that many PPR proteins have sequence-specific RNA binding activity and are involved in organellar gene expression, including RNA transcription, RNA editing, RNA maturation, RNA translation, and RNA splicing (Schmitz-Linneweber and Small, 2008).

In chloroplasts, four PPR proteins have been shown to be involved in RNA splicing of specific introns. PPR4 (Schmitz-Linneweber et al., 2006) was identified in maize and presents, in addition to its PPR motifs, an N-terminal RNA recognition motif (RRM). RNA immunoprecipitation showed that PPR4 was associated with the first intron of rps12 pre-mRNA. In addition, ppr4 mutants are defective in splicing of this intron, demonstrating a direct role of PPR4 in trans-splicing of rps12 intron 1. The Arabidopsis PPR protein OTP51 (for ORGANELLAR TRANSCRIPT PROCESSING 51) carries two C-terminal LAGLIDADG motifs in addition to seven canonical PPR motifs (de Longevialle et al., 2008). LAGLIDADG homing endonucleases are encoded by mobile genetic elements found mostly in group I introns and catalyze DNA cleavage at specific sites lacking the endonuclease catalytic domain (Chevalier and Stoddard, 2001). Most of the LAGLIDADG proteins are endonucleases, but some of them have acquired a dual function, acting both as homing endonucleases and as RNA splicing maturases. OTP51 has lost the amino acids thought to be required for endonuclease activity, but has conserved those described to be essential for the maturase activity. Reverse genetics has shown that OTP51 is involved in RNA splicing of ycf3 intron 2 and, in addition, influences the splicing of a subset of tRNA introns and the atpF intron (de Longevialle et al., 2008). Two other chloroplast PPR proteins may be implicated in splicing, although the data suggest that the effects on splicing may be indirect consequences of alterations in RNA stability. Arabidopsis hcf152 mutants present reduced amounts of spliced petB RNA and are affected in the accumulation of transcripts cleaved between the genes psbH and petB (Meierhoff et al., 2003). HCF152 may bind the exon–intron junction of petB and be directly involved in splicing of the petB transcript (Nakamura et al., 2003) or, alternatively, it may play an indirect role by protecting mature mRNAs from degradation, as shown for PPR10 in maize (Pfalz et al., 2009). Another possible splicing factor is PPR5, a maize P-class PPR protein that stabilizes unspliced trnG-UCC precursor but may also promote its splicing (Beick et al., 2008; Cushing et al., 2005).

In mitochondria, PPR proteins are also likely to be involved in RNA splicing (Figure 3). Most mitochondrial introns are
localized in nad genes, suggesting that splicing mutants are likely to be affected in mitochondrial complex I, leading to partial sterility and delayed development. OTP43 is so far the only PPR protein characterized to be involved in splicing in plant mitochondria (de Longevialle et al., 2007). This P-class PPR is required for trans splicing of nad1 intron 1 and the splicing defect leads to a complete loss of Complex I activity. The otp43 mutant exhibits increased expression of alternative NAD(P)H dehydrogenases and alternative oxidases to bypass the Complex I defect, similar to that seen in rotenone-treated cells (Garmier et al., 2008), the ndufs4 mutant in Arabidopsis (Meyer et al., 2009), and the Nicotiana CMSII mutant (Dutilleul et al., 2003).

All five PPR proteins identified to be involved in splicing in organelles so far are P-class proteins, as opposed to the large class of PLS-class PPR proteins involved in RNA editing (Cai et al., 2009; Hammani et al., 2009; Kim et al., 2009; Okuda et al., 2009, 2010; Robbins et al., 2009; Schmitz-Linneweber and Small, 2008; Takenaka, 2010; Tseng et al., 2010; Verbitskiy et al., 2009; Yu et al., 2009; Zehrmann et al., 2009). The scarce data on the RNA binding activity of P-class PPR proteins suggest they generally bind tightly and specifically to their RNA targets (Schmitz-Linneweber and Small, 2008) and all appear to be primarily required for splicing of a single, specific intron.

**OTHER SPLICING FACTORS**

CRS2 (for Chloroplast RNA Splicing 2) is a peptidyl-tRNA hydrolase (pth\(^{19}\)) that was first described from maize chloroplasts (Jenkins and Barkan, 2001). This protein, in association with the CRM-containing proteins CAF1 and CAF2, is required for the splicing of nine group II introns (rps12 intron 1, ndhB intron, petB intron, ndhA intron, ycf3 intron 1, rpl16 intron, rps16 intron, petD intron, and trnG intron) in maize (Jenkins and Barkan, 2001; Jenkins et al., 1997) (see paragraph on CRM proteins). In Arabidopsis, in addition to a plastid-targeted protein, three orthologs of CRS2 have been described. Two of them are predicted to be targeted to mitochondria whilst the last does not appear to contain a predictable targeting signal (Table 1). The lack of several conserved amino acids important for the activity of the E. coli pth\(^{19}\) enzyme, an enzyme that cleaves the ester bond linking the tRNA and nascent peptide (Menninger, 1976), suggests that CRS2 and its three orthologs may lack peptidyl-tRNA hydrolase activity but have acquired a function in group II intron splicing. Indeed, in the case of CRS2, no function in translation has been observed (Jenkins and Barkan, 2001). However, interestingly, one mitochondrial paralog of CRS2 in Arabidopsis (At5g19830) has conserved PTH-like traits, suggesting that this particular protein could...
have a dual function in RNA splicing and translation (Ostheimer et al., 2005). In Arabidopsis, no mutants in this family have been characterized. However, Arabidopsis CR52 homologs might have the same functions in splicing of several group II introns in association with AtCAF1 and AtCAF2 proteins. In maize, it was impossible to reconstitute splicing in vitro with these three factors together, suggesting that additional proteins are required. One of these additional factors might be RNC1, which has been identified by mass-spectrometry analysis of proteins that co-immunoprecipitate with CAF1 and CAF2 (Watkins et al., 2007).

RNC1 is a plant-specific protein with two ribonuclease III domains that binds RNA but lacks endonuclease activity. In bacteria, Rnase III functions in the processing of a polycistrionic rRNA precursor. In maize, rnc1 mutants are deficient in plastid ribosome accumulation; however, it was not possible to prove a physical association between RNA sequences and RNC1. This, in addition to the lack of endonuclease activity of RNC1, suggests that in plants, RNC1 is not involved in processing polycistrionic RNA precursors but that the phenotype in ribosome accumulation might be a secondary effect of the rnc1 mutation. It has been shown that RNC1 is associated with a subset of group IIA (trnV intron, trnK intron, trnI intron, trnA intron, atpF intron, and rps12 intron 2) and IIB introns (ndhB intron, petB intron, petD intron, trnM intron) and facilitates their splicing (Watkins et al., 2007). Interestingly, RNC1 is also associated with ycf3 intron 1 and rpl2 intron but seems not to be required for their splicing. In addition, this reverse genetic approach has shown that whereas RNC1 was first identified by its co-immunoprecipitation with CAF proteins, it does not seem to be essential for introns that require both CAF/CRS2 complexes, suggesting it plays a supporting role.

Another protein identified by mass-spectrometry analysis of proteins that co-immunoprecipitate with CAF1 and CAF2 contains a domain of unknown function (DUF860), renamed recently as the PORR (for Plant Organelle RNA Recognition) domain due to its ability to bind RNA (Kroeger et al., 2009). WTF1 (for ‘What’s This Factor?’) is part of a family that has 17 members in rice and 15 in Arabidopsis and that comprises 14 orthologous groups. Nearly all these members are predicted to be targeted to chloroplasts or mitochondria. WTF1 is a plant-specific protein, which co-immunoprecipitates from chloroplast extracts with group II introns and is involved in their splicing (Kroeger et al., 2009). This new factor, by forming a heterodimer with RNC1, enhances the splicing of all the RNC1-specific introns. The many other members of this family of organellar proteins have not been characterized so far. Like WTF1, they might be involved in organellar RNA metabolism, opening a new research area of RNA binding proteins, and putative RNA splicing factors.

In chloroplasts, ZmWHY, which has been identified in maize, could be also considered as a splicing factor (Prikryl et al., 2008). ZmWHY is a member of the plant-specific ‘Whirly’ protein family, whose members have been described as DNA-binding proteins that influence nuclear transcription and telomere maintenance and bind nucleoids in chloroplasts and mitochondria. ZmWHY promotes atpF intron splicing in association with CRS1. Many other WHY proteins exist in plants, raising the question of their functions in RNA splicing.

In mitochondria (Figure 3), two additional splicing factors have been described that do not belong to the different families of proteins described above. PMH2 (for Putative Mitochondrial RNA Helicase 2) is a DEAD-box protein (Kohler et al., 2009; Matthes et al., 2007). In Arabidopsis, 58 genes encoding DEAD-box proteins have been identified (Boudet et al., 2001; Mingam et al., 2004). Through their primary function in modifying inter- or intramolecular RNA structures or in dissolving RNA protein complexes, DEAD-box proteins are potentially involved in most events of RNA processing, including RNA synthesis, RNA modification, RNA cleavage, RNA degradation, as well as ribosome biogenesis and translation initiation (Linder, 2006). PMH2 is associated with large RNA-containing complexes in mitochondria (Matthes et al., 2007) and is required for the efficient splicing of 15 out of the 23 group II introns found in Arabidopsis mitochondria (Kohler et al., 2009). More precisely, the pmh2 mutant presents a decrease in the splicing of: nad1 intron 2 and 3; nad2 intron 1, 2, and 4; nad4 intron 2 and 3; nad5 intron 1, 2, and 3; nad7 intron 1 and 4; cox2 intron; rps3 intron; and rpl2 intron. As suggested by the large number of introns that are more or less affected in the mutant, PMH2 does not seem to have a preference for a particular substrate and rather functions as a RNA chaperone that influences RNA stability. Three other DEAD-box proteins involved in RNA splicing have been characterized in other species. In S. cerevisiae, Mss116 is important for the splicing of all mitochondrial group I and group II introns (Huang et al., 2005) and Ded1 has been shown to facilitate group II intron splicing in vitro (Solem et al., 2006). In N. crassa, CYT-19 enhances group I intron splicing (Bertrand et al., 1982).

**ROLE OF SPICING FACTORS IN RNA FOLDING AND SPICE SITE RECOGNITION**

A large set of splicing factors has been identified in recent years, mainly by reverse genetic approaches and/or because of their RNA binding activities. However, their precise roles and the basis for sequence-specific RNA recognition still need to be elucidated for most of them. Progress in understanding splicing mechanisms has been achieved for some chloroplast splicing factors, providing new insights into ribonucleoprotein complexes and the need for RNA folding trans-factors.

As described above, the CRS1 protein, which has three CRM domains, binds the atpF intron in vivo (Ostheimer et al., 2003; Till et al., 2001). To further understand its role in RNA splicing, a recombinant dimeric CRS1 has been used for binding assays on the atpF intron (Ostertset et al., 2005). It revealed that the maize recombinant CRS1 binds two different sequences within
the maize atpF intron, which correspond to 111 nucleotides at the 3’ end of domain I and 30 nucleotides (680–710) in domain IV. Through this association, CRS1 promotes the RNA folding of the atpF intron into its predicted ‘catalytically active’ form. Interestingly, these sequences are not well conserved between monocotyledons and dicotyledons. For example, a deletion of the 3’ end of domain I is observed in dicotyledons. Because the Arabidopsis homolog of CRS1 was shown to be required for the splicing of the same intron (Asakura and Barkan, 2006), it would be very interesting to discover the sites bound by the Arabidopsis protein to understand how CRS1 recognizes its binding sites and how RNA binding proteins have co-evolved with their RNA targets. To understand how multiple CRM domains cooperate to achieve highly specific RNA binding, affinity assays were performed on the three CRM domains of maize CRS1 (Keren et al., 2008). Through this analysis, these three CRM motifs did not show the same affinity for the RNA sequences of the atpF intron. Whereas CRM domain 3 bound tightly to atpF intron, CRM domain 1 associated with the 111 nucleotides at the 3’ end of domain I and CRM domain 2 interacted with several regions within domains I, III, and IV but with low affinity.

A similar approach was used to investigate the binding sites of PPR5 (Beick et al., 2008; Williams-Carrier et al., 2008). ZmPPR5 binds in vivo to the unspliced precursor of trnG-UCC and is involved in RNA stabilization and possibly RNA splicing of this intron. Recombinant ZmPPR5 is associated with high affinity to the central region of the trnG intron (Beick et al., 2008) and interacts preferentially with single-stranded RNA in a region of 50 nucleotides in domain I of the trnG intron (Williams-Carrier et al., 2008). This binding site is located where the intron is cut in the mutant, confirming a role for PPR5 in RNase protection.

Through these analyses and RIP-Chip experiments, it seems quite clear that at least some splicing factors directly interact with introns in a sequence-specific manner. By this binding activity, the interacting protein changes the intron RNA folding to enhance its splicing. This conformational change is mediated by at least one or more proteins, depending on the intron targeted. For example, even if we cannot exclude that many splicing factors remain to be identified, only one splicing factor has been identified for ycf3 intron 2, clpP intron 2, or trnl intron, while the splicing of some other introns involve four or five known splicing factors, such as, for example, the petB intron. In the case of the atpF intron, six splicing factors have been described to be absolutely or partially required for its splicing. Sedimentation through sucrose gradients showed that CRS1 is part of a ribonucleoprotein complex of about 600–700 kDa (Till et al., 2001), comprising about 270 kDa of the atpF intron and around 160 kDa of dimeric CRS1. By adding in the molecular weight of all the splicing factors (WTF1, 61 kDa; RNC1, 62 kDa; MatK, 65 kDa; ZmWHY, 29 kDa) that have been shown to be involved in atpF splicing, a complex of 647 kDa, very close to the experimental observations, is predicted. This raises the question of whether all the splicing factors are present in a single RNP complex. Splicing factors may interact with each other via direct protein–protein interactions or indirectly by RNA tethering. For example, whilst RNC1 and WTF1 were shown to interact together to promote the splicing of most group II introns in chloroplasts (Kroeger et al., 2009), ZmWHY is associated with CRS1 in an indirect RNA-mediated manner to promote the splicing of atpF (Prikryl et al., 2008). So far, the evidence suggests that factors required for splicing more than one intron are generally incorporated into one of several alternate multi-protein complexes (Kroeger et al., 2009). Intron-specific factors have not generally been found to be physically associated with such splicing complexes. It is possible that they act in an initial sequence-specific binding step that alters RNA conformation ready for recruitment of the relevant splicing complex, but such ideas await the challenging reconstitution of RNA splicing in vitro before they can be tested.

OTHER MECHANISMS INFLUENCING RNA SPlicing EFFICIENCY INDIRECTLY

RNA Editing

Maturation of organellar mRNAs involves many different processes in plant organelles, which include exonucleolytic processing of 5’ and 3’ termini, RNA splicing, and RNA editing. RNA editing is usually an early step of RNA maturation and required for expression of many plant mitochondrial and chloroplast genes. Thirty-four cytidine to uridine editing sites have been described in Arabidopsis chloroplast transcripts (Chateigner-Boutin and Small, 2007) whereas 441 were found in Arabidopsis mitochondria, 427 in Brassica napus, and 491 in Oryza sativa (Giege and Brennicke, 1999; Handa, 2003; Notsu et al., 2002). In addition to their principal effect of changing coding sequences of mRNA, RNA editing sites have also been found in introns or near intron–exon junctions. The function of RNA editing sites in untranslated regions remains to be elucidated, but a role in intron RNA folding is plausible. Such effects were observed in a chimeric yeast mitochondrial intron containing the domain V of nad1 intron 3 from Oenothera mitochondria. In these experiments, in vitro self-splicing was observed only for the edited and not for the unedited mRNA (Borner et al., 1995). In addition, some RNA editing events close to exon–intron boundaries are affected by splicing: only unspliced or spliced RNA is edited, depending on the mRNA. For example, in the case of cox2 exon 1, suppressing an editing site located in the IBS1 at low temperature reveals an accumulation of unspliced precursors in wheat, suggesting that this editing event is required for high splicing efficiency (Kurihara-Yonemoto and Handa, 2001). On the other hand, some RNA editing sites very close to intron–exon junctions are edited only in the spliced mRNA, suggesting that the recognition of the editing site occurs only when the precursor is spliced or that the different EBS and IBS interactions are
competing with the editing machinery (Li-Pook-Than and Bonen, 2006).

Temperature
RNA folding is sensitive to temperature. Interactions within introns or between intron and exon sequences may be different at low or high temperatures, leading to different RNA conformations. Changes in RNA folding will modify the interactions between introns and their associated splicing factors and thus modify splicing efficiency. Such alterations have been observed for ndhB mRNA in tobacco (Karcher and Bock, 2002). The processing of ndhB pre mRNA is sensitive to heat stress. At 25 or 37°C, the splicing of this intron is normal, but no spliced ndhB is observed at 42°C and an accumulation of the unspliced mRNA is observed. In this case, the splicing defect may be explained by an inactivation of the associated splicing factors at high temperatures. Another example of such an event has been described for ycf3 intron 1 in the cl3 mutant (for cytoplasmic line 3) in barley (Landau et al., 2009). Two point mutations in ycf3 intron 1 lead to a decrease in splicing of this intron at 18°C and no splicing was detected when plants were grown at 32°C. In silico folding analysis of the mutant ycf3 intron 1 sequence suggested that the temperature increase to 32°C might lead to the disruption of the RNA secondary structure required for intron splicing.

CONTROL OF ORGANELLAR RNA SPlicing BY SPlicing FACTORS: A HYPOTHETICAL MODEL

Most of the nuclearly encoded splicing factors that have been identified are required for RNA splicing of their corresponding introns. If splicing were rate-limiting for expression of the organelle proteins encoded by the spliced transcripts, then these splicing factors could be said to regulate organelar gene expression. A hypothetical model of how this might occur is shown in Figure 4. The model suggests that splicing factors can be classified into two types: sequence-specific splicing factors and general splicing factors. Sequence-specific splicing factors recognize intron sequences and initiate splicing by recruiting general splicing factors. General splicing factors then catalyze the splicing reaction.

Figure 4. Hypothetical Model of Plant Organelar RNA Splicing.
(A) A primary transcript is produced that is incapable of self-splicing under physiological conditions.
(B) One or more primary-sequence-specific RNA binding factors, such as a PPR protein for example, associates with the selected intron. This initial protein–RNA association leads to conformational change stabilizing the RNA in a structure that can be recognized by general splicing factors. In the case of trans-splicing introns, this RNA binding factor could associate with the two parts of the intron, helping in this way the RNA–RNA interaction.
(C) Association of one of several alternative general splicing complexes enhances the two trans-esterification steps leading to splicing. Two predictions from this hypothetical schema can be made: the first is that there is no requirement that the initial sequence-specific splicing factors targeting the selected intron have any physical association with the general splicing complex. The second is that the binding of the initial RNA binding factor is likely to be the rate-limiting step, making these intron-specific splicing factors good candidates as regulatory factors.
expression, as changes in the level of activity of the splicing factor should lead to concomitant changes in the level of spliced transcripts and their translation. It remains unclear whether any such regulation occurs. This question was partially addressed in a study of mitochondrial transcript profiles in wheat during early germination and seedling development (Li-Pook-Than et al., 2004). During early development, RNA splicing was shown to be incomplete for many introns. For example, in the case of nad7 mRNA, partially spliced mRNA accumulated during the early developmental stages before subsequently disappearing 4 d after germination. Similar observations were made for nad4 mRNA, where nad4 intron 1 was inefficiently excised during early development, whilst introns 2 and 3 were fully spliced. In chloroplasts, examination of four genes in several tissues of maize suggests that splicing is part of the regulation of plastid gene expression (Barkan, 1989). While the fraction of spliced transcripts of atpF, petB, petD, and rpl16 is similar in all leaf tissues examined, it is lower in meristematic and root tissue. Given that there is little evidence for gene-specific transcriptional regulation in plant organelles and many assumptions that regulation of gene expression is largely post-transcriptional, intron-specific splicing factors make attractive candidates as potential regulatory factors (see the model presented in Figure 4). It would be useful to carry out a systematic analysis of rates of splicing and expression of splicing factors throughout plant development and correlate the observations with translation rates for organelar proteins. One such analysis has been partially done in rnc1 mutants. Three alleles have been identified in maize for which two are mild rnc1 alleles. While some introns are less affected in the mild rnc1 alleles than with a total loss of function of RNC1, there is a strong effect on trnA and trnI splicing, indicating that RNC1 is limiting for splicing in these cases and that it could be a regulatory factor (Watkins et al., 2007).

CONCLUSIONS AND FUTURE PROSPECTS

RNA splicing in plant organelles is considerably more complicated that was first expected, involving a large number of diverse RNA binding factors. A wide range of nuclearly encoded proteins have co-evolved with introns, giving rise to several novel plant-specific RNA binding protein families. Some splicing factors are involved in many intron splicing events. While no data are available concerning the recognition sites of these factors, it could be proposed that these proteins bind to specific sequences that are common to all targeted introns or, more likely, that they recognize secondary structures specific to each intron class. Other splicing factors are specific to only one intron, suggesting a specific RNA recognition site, which may be determined by primary sequence or unique structural motifs. The result is that each intron requires a different set of splicing factors. In theory at least, this provides the necessary flexibility to control expression of any spliced transcript in a gene-specific manner. It remains to be seen whether regulated activity of individual splicing factors is a factor in controlling any aspects of organelle gene expression, but the possibility is there. It will be interesting in the near future to understand how these splicing factors are regulated during plant growth and development, and also to see their responses to imposed environmental stress. From a mechanistic point of view, more splicing factors need to be characterized to catalog the splicing machinery for each intron, with the ultimate goal of examining the splicing mechanism in vitro with a reconstituted splicing complex.

FUNDING

A.F.L. was supported by a fellowship from the AGRON-OMICS Integrated Project (LSHG-CT-2006-037704). I.D.S. is supported as a WA State Premier’s Fellow. No conflict of interest declared.

REFERENCES


family genes in *Arabidopsis*, *Caenorhabditis*, and Drosophila. Genome Res. 11, 2101–2114.


Till, B., Schmitz-Linneweber, C., Williams-Carrier, R., and Barkan, A. (2001). CRS1 is a novel group II intron splicing factor that was derived from a domain of ancient origin. RNA. 7, 1227–1238.


