MEETING REPORT

The blossoming of RNA biology: Novel insights from plant systems

JÉRÔME BOVE,1 CAREY L.H. HORD,2,3 and MELISSA A. MULLEN4
1Department of Biology, The Pennsylvania State University, University Park, Pennsylvania 16802, USA
2Department of Biology and the Huck Institute for the Life Sciences, The Pennsylvania State University, University Park, Pennsylvania 16802, USA
3Integrative Biosciences Graduate Degree Program, The Pennsylvania State University, University Park, Pennsylvania 16802, USA
4Department of Chemistry, The Pennsylvania State University, University Park, Pennsylvania 16802, USA

INTRODUCTION

Scientists from around the world met in May 2006 at the Sixteenth Penn State Symposium in Plant Physiology to discuss insights into RNA biology from both plant and nonplant systems. This meeting was sponsored in part by the National Science Foundation and the US Department of Agriculture, and included scientists at the forefront of this exciting field. Experts with diverse backgrounds, from plant molecular biologists to structural biochemists, discussed new advances in RNA biology. The topics presented covered the biological roles of RNA, particularly hormonal responses and development, regulation of gene expression by RNA, regulation of RNA localization, and bioinformatics approaches/analysis tools. This article will give an overview of small RNA pathways in plants, the biological roles of RNAs, and methods for analysis that were discussed in this meeting.

MECHANISMS AND MAJOR PATHWAYS OF SMALL RNAs

Small RNAs (sRNAs) are numerous in sequence, widely distributed among eukaryotes (particularly plants), and affect many biological processes. sRNAs include microRNAs (miRNAs), small interfering RNAs (siRNAs), trans-acting siRNAs (ta-siRNAs), and natural antisense transcript siRNAs (nat-siRNAs). Each type of sRNA is unique with respect to length, function, and the plant Dicer-like (DCL) endonucleases that are involved in its pathway. In general, siRNAs are 21–24 nucleotides (nt) in length and their biogenesis involves DCL2, DCL3, and DCL4. ta-siRNAs are only 21 nt long and involve only DCL4, while the 21 and 24 nt nat-siRNAs involve DCL1 and DCL2, respectively. In addition, miRNAs are typically 21 nt in length, and their biogenesis involves DCL1 (Jones-Rhoades et al. 2006).

miRNAs are sRNAs that often direct silencing of target mRNA(s), which are distinct from the miRNA locus itself (Jones-Rhoades et al. 2006). The transcription product of a miRNA gene, called the pri-miRNA (Xie et al. 2005), contains an imperfect inverted repeat sequence that is separated by a span of RNA that allows the strand to fold back on itself and form a stem–loop structure. The pri-miRNA is cleaved on either side of the miRNA sequence by the enzyme Dicer (DCL1 in Arabidopsis) forming what is called the miRNA–miRNA* duplex (Park et al. 2002; Reinhart et al. 2002; Lim et al. 2003; Kurihara and Watanabe 2004; Jones-Rhoades et al. 2006). In addition to DCL1, the Arabidopsis HYONASTIC LEAVES1 (HYL1) RNA-binding protein, the identification and characterization of which Nina Fedoroff (Penn State University) described at the meeting, is required for the cleavage of pri-miRNA and DCL1 and HYL1 likely function together in this process (Lu and Fedoroff 2000; Jones-Rhoades et al. 2006). Hai Huang (SIPPE, Shanghai, and Joint Center of Penn State-SIPPE-SJITU) discussed the involvement of SERRATE (SE), an Arabidopsis C2H2 zinc-finger protein (Grigg et al. 2005), in this process (Yang et al. 2006b). SE and HYL1 are localized to the nucleus and were shown to physically interact via a yeast two-hybrid experiment as well as a glutathione S-transferase (GST) pull-down assay. Thus, DCL1, HYL1, and SE appear to all function together in pri-miRNA cleavage. After cleavage, the miRNA–miRNA* duplex is methylated on the 3’ hydroxyl group of the 3’ terminal nucleotides. This methylation requires the Arabidopsis HUA ENHANCER1 (HEN1) gene encoding a methyl transferase that can methylate the miRNA–miRNA* duplexes in vitro (Ebhardt et al. 2005; Yu et al. 2005a; Li et al. 2005b). The methylation of the miRNA–miRNA* duplex appears to stabilize the miRNA by protecting it...
DCL2 produces nat-siRNAs, which are required for sense RISC and help to silence genes post-transcriptionally. role in producing sRNAs that comprise part of the antiviral defense. This suggests that plants require at least five DCL proteins (maize and wheat have five, while rice has six). This indicates that 99% of known siRNA clusters are methylated in the promoter region in a tissue-specific manner. Interestingly, pseudogenes are highly methylated, while the miRNA strand is degraded (Reinhart et al. 2002; Jones-Rhoadeas et al. 2006). The key component of the miRNA RISC in Arabidopsis is ARGONAUTE1 (AGO1). Many Argonaute proteins contain both an RNA-binding and an RNA-cleavage domain. A major difference between plants and animals is that in plants the miRNA loaded into the RISC directs cleavage of complementary (target) mRNA strands by AGO1. After cleavage, the mRNA is released and another target mRNA can be loaded onto RISC for cleavage.

Joe Ecker (Salk Institute) observed that ~10,000 genes in Arabidopsis are methylated. Among these genes, about 10% are methylated in the promoter region in a tissue-specific manner. Interestingly, pseudogenes are highly methylated, and their promoter and coding regions tend to be equally methylated. There is also a strong correlation between methylation and heterochromatic regions. Another interesting observation is that 99% of known siRNA clusters are methylated.

The RNAi pathway is used to regulate gene expression and to defend against viral attack. This complex pathway targets long double-stranded RNA and viral RNAs for destruction. The double-stranded RNA is recognized by one of the DCL protein complexes and is cleaved into 21–24 nt pieces of double-stranded RNA. These siRNAs can then be used for short- or long-distance signaling. Some siRNAs can continue on to join the RISC to cleave target RNAs and stop expression of the target gene.

Peter Waterhouse (CSIRO, Canberra, Australia) discussed differences between the DCL proteins across plant species. Where it is known, dicots have at least four DCL proteins (Arabidopsis has four, poplar has five) and monocots have at least five DCLs (maize and wheat have five, while rice has six). This suggests that plants require a minimum of four DCL proteins, and that after the monocot–dicot split, the monocots acquired at least one additional DCL gene (Margis et al. 2006).

James Carrington (Oregon State) talked about the roles of DCL4 and DCL2 in viral defense. DCL4 plays a major role in producing sRNAs that comprise part of the antiviral RISC and help to silence genes post-transcriptionally. DCL2 produces nat-siRNAs, which are required for sense transgene RNA silencing in viral defense. Vicki Vance (University of South Carolina) also discussed the roles of DCL2 and DCL4 in viral defense. Vance has shown that DCL3 is not needed for silencing of endogenous genes by geminivirus-mediated VIGS (DNA-VIGS); however, Niki Robertson (North Carolina State University) suggested that DCL3 may play a role in the silencing of DNA geminiviruses with DCL2 and DCL4. Robertson also described several differences in DNA VIGS response for different ecotypes of Arabidopsis. These differences indicate that the pathways for RNA silencing and viral response may be more variable than previously thought. Carrington also reported the roles of two other proteins in viral silencing responses, RDR6 and RDR1, which aid in antiviral immunity by making dsRNA from single-stranded viral templates (Deleris et al. 2006).

The mechanism of gene silencing by hairpin RNAs (hpRNAs) was detailed by Peter Waterhouse. An interesting feature of gene silencing in plants is that it does not seem to be temporally self-perpetuating without active transcription of the silencing RNA. This property allows the use of specific/inducible promoters to transiently knock down gene expression. In addition, grafting experiments in tobacco showed that systemic long-distance silencing is triggered by hpRNAs, comparable to viral RNA-induced silencing. Studies of Arabidopsis dcl single and double mutants showed that both hpRNA and virus-induced gene silencing (VIGS) involve DCL2, DCL3, and DCL4, particularly DCL2 and DCL4. These results suggest that hpRNAi operates through the VIGS pathway. This hypothesis is further supported by the suppression of hpRNA-mediated RNAi by HC-Pro, a viral protein that suppresses the VIGS pathway in plants (Anandalakshmi et al. 1998). VIGS is an efficient technique used to mimic a loss-of-function mutation without stable transformation. Savithramma Dinesh-Kumar (Yale University) presented intriguing applications of this strategy in studying plant development, viral movement, and plant disease resistance.

SMALL RNAs ARE KEY REGULATORS OF PLANT DEVELOPMENT

Leaf primordia are initiated at the periphery of the shoot apical meristem. For proper leaf development, it is necessary for the leaf primordium to establish polarity within its three-dimensional planes: the mediolateral, proximodistal, and abaxial–adaxial (dorsoventral) polarities (Hudson 2000). The abaxial–adaxial polarity is necessary for establishing the mediolateral polarity. In Arabidopsis, several genes are involved in regulating abaxial–adaxial polarity. FILAMENTOUS FLOWER (FIL), a member of the YABBY family, and YAB3 promote the abaxial leaf fate (Siegfried et al. 1999). The class III HD-ZIP family genes, REVOLUTA (REV), PHABULOSA (PHB), and
PHAVOLUTA (PHV), promote adaxial leaf differentiation (McConnell and Barton 1998; McConnell et al. 2001). In addition, ASYMMETRIC LEAVES1 (AS1), which is an R2–R3 MYB domain protein, and AS2 form a complex and specify adaxial leaf identity by positively regulating REV, PHB, and PHV, and by negatively regulating FIL (Li et al. 2005a).

Kathy Barton (Carnegie Institution of Washington, Stanford) and Hai Huang discussed the role of miRNA and post-transcriptional gene silencing (PTGS) in regulating leaf polarity and development. miR165/166 are complimentary to a portion of the region encoding the START domain of REV, PHB, and PHV and can cause the cleavage of PHB and PHV mRNA in vitro (Emery et al. 2003; Tang et al. 2003; Zhong and Ye 2004; Mallory et al. 2004). Mutations in a miRNA-binding site, which do not alter the amino acid sequence coded by the REV, PHB, or PHV genes, cause expansion of their expression to the abaxial region, leading to adaxialization of leaves and other organs (Emery et al. 2003; Mallory et al. 2004). In addition, promoter–GUS fusions showed polar expression of PHB in the adaxial leaf domain, indicating that mir165/166 represses PHB throughout the developing leaf, and not just in the abaxial domain (Bao et al. 2004). Interestingly, Barton also reported that in the phb-d mutant, regions of the PHB gene had reduced cytosine methylation, suggesting the possibility of miRNA-directed DNA modification (Bao et al. 2004).

Recent work from Hai Huang’s laboratory has revealed a role for RDR6 in leaf development. RDR6 was known previously to be involved in PTGS (Dalmay et al. 2000; Mourrain et al. 2000), and to be important for vegetative phase change (Peragine et al. 2004). This laboratory earlier identified ASYMMETRIC LEAVES1 and 2 (AS1 and AS2), two genes whose products interact together and play a role in the adaxial–abaxial polarity formation (Xu et al. 2003). Disruption of the RDR6 gene in the as2 mutant background causes severely enhanced developmental defects, including higher frequencies of abaxialized leaves (Li et al. 2005a). This correlated with a decrease in PHB and REV expression, expanded FIL expression, and an increase in miR165/166 levels. Thus, the AS1–AS2 pathway and the RDR6 pathway both repress MIR165/166. In the absence of the AS1–AS2 and RDR6 pathways, abnormally high levels of miR165/166 repress expression of the adaxial-promoting factors, leading to abaxialized leaves. However, additional work with ago1 mutants revealed that miR165/166 are not the only determinants of leaf polarity (Yang et al. 2006a). Furthermore, more recent work has shown that RDR6, SGS3, and AGO7 function in the same pathway to direct leaf polarity (Adenot et al. 2006; Fahlgren et al. 2006; Garcia et al. 2006; Xu et al. 2006).

Isabel Bäurle from Caroline Dean’s laboratory (John Innes Center, UK) and Josef Kuhn from Julian Schroeder’s laboratory (UC San Diego) discussed examples of how different genes involved in RNA processing affect flowering time. The transition from vegetative growth to flowering is governed by environmental conditions, particularly day length and extended periods of low temperature. A central figure in this pathway is FLOWERING LOCUS C (FLC), which encodes a MADS-domain transcription factor that acts as the main repressor of flowering (Michaels and Amasino 1999; Simpson and Dean 2002). FLC is regulated by many factors; Isabel Bäurle and Josef Kuhn discussed those that post-transcriptionally regulate FLC expression: ABA HYPERSENSITIVE1 (ABH1), FCA, and FY (Macknight et al. 1997; Simpson et al. 2003; Bezerra et al. 2004).

ABH1 encodes the large subunit of an Arabidopsis nuclear mRNA cap binding protein (CBP80) (Hugouvieux et al. 2001). In the abh1 mutant, FLC and FLOWERING LOCUS M (FLM) expression are decreased. abh1 specifically affects intron 1-dependent splicing of downstream introns in FLC pre-mRNAs and also reveals differential premature polyadenylation within intron 1 of FLM transcripts compared to wild type. The reduction of both FLC and FLM transcript levels provides a basis for early flowering in abh1. In addition to regulating FLC and FLM expression, Kuhn reported that ABH1 also affects CONSTANS (CO) expression. CO is a pivotal regulator in the photoperiodic pathway that promotes flowering, and CO transcripts are affected in abh1 (J. Kuhn and J. Schroeder, unpubl.).

FCA encodes an RNA-recognition motif (RRM) nuclear RNA-binding protein that is also involved in pre-mRNA processing (Macknight et al. 1997; Macknight et al. 2002). FY is an RNA 3′-end processing factor that interacts with FCA (Quesada et al. 2003; Simpson et al. 2003). Bäurle reported that together FCA and FY cause silencing of FLC via the siRNA pathway. Interestingly, FCA itself is also regulated post-transcriptionally through alternative splicing and polyadenylation of its introns (Macknight et al. 1997; Macknight et al. 2002; Quesada et al. 2003). Of the four FCA transcripts produced, α, β, γ, and δ, only the γ form is functional. Together with FY, FCA regulates RNA processing of its own pre-mRNA via polyadenylation and thereby negatively regulates its own expression (Quesada et al. 2003; Simpson et al. 2003). FY likely has other essential functions beyond controlling flowering time as a null fy mutant is embryo lethal (Henderson et al. 2005).

In addition of exploring the role of RNA processing in flowering time, the role of small RNAs in flower organ development was also discussed. Flowers typically have four types of organs: sepals (outer covering), petals (attractive to insects), stamens (male), and carpels (female). The classic ABC model for determining floral organ identifies states that the A function promotes the identity of sepals, combined A and B functions specify petals, combined B and C functions designate the stamens, and the C function directs the carpel identity (Coen and Meyerowitz 1991; Weigel and Meyerowitz 1994; Ma 2005). Disruption of the ABC genes causes homeotic mutations. For example,
disruption of the Arabidopsis C function gene AGAMOUS (AG) results in flowers with alternating whorls of sepals and petals, but no reproductive organs, namely the stamens and carpels (Yanofsky et al. 1990). This is partly due to the mutually repressive nature of the A and C function genes (Drews et al. 1991). AG represses expression of the A function gene APETELA1 (AP1), while AP2 negatively regulates AG expression. The HUA and HUA ENHANCER (HEN) genes contribute to the C function in flower development (Chen and Meyerowitz 1999; Chen et al. 2002; Western et al. 2002). Several HUA and HEN genes code for proteins that may be involved in RNA processing (Cheng et al. 2003).

Xuemei Chen (UC Riverside) discussed elegant research from her laboratory and others' laboratories that is beginning to uncover how the miRNA pathway is involved in floral organ identity determination and in controlling floral organ number. Disruption of either HEN1 or DCL1 results in stamen-to-petal transformation along with a decrease in miRNA abundance, indicating that miRNAs negatively regulate A function genes (Chen et al. 2002; Park et al. 2002). As discussed earlier, HEN1 is a sequence-nonspecific sRNA methyltransferase that is required for the maturation of miRNAs, including miR172 (Chen 2004). Plants expressing MiR172 driven by the cauliflower mosaic virus 35S promoter produce flowers phenotypically similar to those of the ap2 mutant. In addition, AP2 mRNA contains a miR172 target site that when mutated (AP2m3) results in flowers with multiple whorls of stamens. Together these data indicate that miR172 is a negative regulator of AP2. In addition, the extra stamen phenotype seen in AP2m3 plants implies that AP2 positively regulates stem cell propagation. WUSCHEL (WUS) is a positive regulator of stem cell proliferation, and in the wus mutant background the extra stamen phenotype of the AP2m3 plants was no longer observed (Wurschum et al. 2006). Further work showed that AP2 is a positive regulator of WUS, thereby promoting stem cell proliferation and floral organ number. miR172 also restricts the expression of both AP3 and PI (B function genes), while AP2 is a positive regulator of their expression.

**SMALL RNAs ARE INVOLVED IN PLANT STRESS RESPONSES**

sRNAs are also involved in plant stress response pathways. Many different environmental conditions can cause stress to plants, including nutrient stress, high salt, and various factors that lead to the generation of reactive oxygen species (ROS). Tzyy-Jen Chiou (Academia Sinica, Taiwan) described the role of miR399 in phosphate homeostasis in Arabidopsis. miR399 accumulates in roots and shoots during phosphate starvation (Fujii et al. 2005; Chiou et al. 2006). Overexpression of miR399 increases phosphate uptake by the roots and its subsequent translocation to the shoot, which eventually causes phosphate toxicity (Fujii et al. 2005; Chiou et al. 2006). miR399 targets an mRNA encoding a ubiquitin-conjugating E2 enzyme, UBC24. The level of the UBC24 mRNA decreases in the roots of phosphate-starved plants, and a loss-of-function ubc24 mutant phenocopies the MIR399 overexpression plants. Interestingly, a previously isolated EMS-induced mutant (pho2) defective in phosphate homeostasis (Delhaize and Randall 1995) was recently shown to carry a nonsense mutation in UBC24 (Aung et al. 2006; Bari et al. 2006).

A compelling example of the role of sRNAs in abiotic stress response was presented by Jian-Kang Zhu (University of California, Riverside), whose laboratory investigates a class of nat-siRNAs (Borsani et al. 2005). Zhu discussed the role of two genes in salt stress, PSCDH and SRO5. In response to salt stress, SRO5 gene expression is induced while PSCDH expression decreases. PSCDH encodes a Δ1-pyrroline-5-carboxylate dehydrogenase enzyme that catalyzes Δ1-pyrroline-5-carboxylate (P5C). This reaction is part of proline catabolism, and a decrease in PSCDH levels is then expected to lead to an accumulation of both proline and P5C. While proline accumulation is a beneficial component of plant salt stress response, accumulation of P5C is toxic (Hellmann et al. 2000; Deuschle et al. 2004), possibly by increasing the production of ROS. SRO5 encodes a mitochondrial protein (Borsani et al. 2005) hypothesized to counteract the side effects of ROS accumulation. SRO5 and PSCDH genes overlap on opposing DNA strands and can produce convergent transcripts whose last 760 nt (before the polyA cleavage site) constitute natural antisense strands. The striking point is that 24- and 21-nt siRNAs that match this region have been cloned from salt-stressed plants (Sunkar and Zhu 2004; Borsani et al. 2005). By monitoring the levels of these nat-siRNAs in many siRNA biogenesis pathway mutants, Borsani et al. (2005) proposed a model for the regulation of PSCDH in response to salt stress with the following steps: (1) the expression of SRO5; (2) the formation of a dsRNA by annealing of the two corresponding natural antisense transcripts between SRO5 and PSCDH; (3) from the dsRNA, a 24-nt nat-siRNA is generated via a DCL2-dependent process; (4) this 24-nt nat-siRNA guides the additional phased and DCL1-dependent cleavages of the dsRNA into 21-nt nat-siRNAs; and (5) the different nat-siRNAs generated then target the PSCDH transcripts for degradation.

According to sequence analysis, about 10% of the genes in the Arabidopsis genome are overlapping, and >20% in the human and fly genomes overlap. Hence, there is an important potential for gene expression regulation through nat-siRNA pathways. Zhu reported that 24-nt nat-siRNAs have been observed for several pairs of overlapping genes, whose roles in stress response are hypothesized to be antagonistic. The high rates of gene overlap in genomes in which genome size does not seem to be a constraint for such arrangements is all the more intriguing.

2038 RNA, Vol. 12, No. 12
Involvement of mRNA processing factors in responses to stresses was also highlighted through studies of the polyadenylation/splicing factor AtCPSF30, and the RNA-binding protein SCI2. Q. Quinn Li (Miami University, Ohio) studies AtCPSF30, which appears to be involved in the response to oxidative stress. A mutant deficient in this factor, named oxt6, has an increased tolerance to oxidative stress and an altered mRNA accumulation profile. Some of the affected mRNAs correspond to oxidative defense genes. AtCPSF30 was shown to be a nuclear RNA-binding protein capable of homodimerization in vitro (Delaney et al. 2006). Interestingly, the RNA-binding activity of AtCPSF30 is inhibited by the binding of calmodulin (Delaney et al. 2006).

Several plant stress–response pathways are mediated by the plant hormone jasmonate. Jasmonate is an important hormone in plant responses to wounding, drought, insects, pathogens, and ozone. Fran Robson from John Turner’s laboratory (University of East Anglia, UK) presented her work on the *Arabidopsis* sci2 mutant, which has a missense mutation in a KH domain RNA-binding protein. This mutant was identified by a screen based on the restoration of VSP gene expression (a jasmonate inducible gene) in the coi-16 mutant, which is defective in the jasmonate perception/response pathway (Benedetti et al. 1995). The sci2 single mutant confers hypersensitivity to jasmonate and wounding, an altered response to red light and resistance to drought, infestation by aphids, and the powdery mildew *Erysiphe cichoracearum*. Gene expression data available through Genevestigator (Zimmermann et al. 2004) indicate that SCI2 expression is transiently induced by jasmonate, wounding, drought, osmotic stress, and salinity. These data, taken together with the RNA-binding nature of SCI2, suggest that SCI2 regulates responses to jasmonate and different stresses through post-transcriptional regulation.

**RNA PROCESSING**

Subcellular localization and compartmentalization of mRNA can be crucial for the correct localization of the corresponding encoded protein, and dramatically affects multiple processes, particularly development (Fedoroff 2002; Crofts et al. 2004; Van de Bor and Davis 2004). A beautiful example was provided by Thomas Okita, from Washington State University. He described how the localization of the main storage proteins in rice seeds, prolamine and glutelin, is controlled through the differential targeting of their mRNAs to two subdomains of the cortical endoplasmic reticulum (ER). Prolamine mRNAs are directed to the intracisternal prolamine protein bodies (PB-ER), a specialized area of the ER where prolamine accumulates and assembles. Glutelin mRNAs are directed to the cisternal ER, and after translation, glutelin protein is translocated to the ER lumen and exported to protein storage vacuoles via the Golgi (Crofts et al. 2004, 2005). To investigate how the mRNAs are targeted to their respective locations, Hamada et al. (2003b) used synthetic constructs containing a GFP gene fused to different portions of the prolamine or glutelin coding and noncoding sequences. These experiments showed that prolamine mRNAs targeted to the PB-ER require two cis-elements (dubbed zicodes): one located in the coding sequence, and a second residing in the 3' UTR. Interestingly, addition of the glutelin 3' untranslated region to the two prolamine zicodes redirects prolamine mRNA to the cisternal ER. Therefore, the targeting of glutelin mRNA to cisternal ER is controlled by a cis-element and does not simply result from default targeting of the mRNAs to the cisternal ER. The importance of differential mRNA targeting to the final localization of these storage proteins was illustrated by the fact that the targeting glutelin mRNAs to the PB-ER disrupts glutelin packaging into protein storage vacuoles. To study the transport of prolamine mRNAs, Hamada et al. (2003a) used a two-gene system consisting of GFP fused to the viral RNA-binding protein MS2 and a hybrid prolamine gene containing tandem MS2 RNA-binding sites. They observed that the prolamine RNAs move as particles, generally unidirectionally at 0.3–0.4 μm/sec, and this transport is dependent on microfilaments.

In addition, mRNA subcellular localization is an important mechanism for regulating gene function. For example, in *Drosophila*, the MAGO NASHI (Mago) protein, which is essential for oocyte polarity, binds with part of the exon–junction complex (EJC), and affects localization of target mRNAs. HAPLESS1 (*HAP1/AtMago*) is an *Arabidopsis* ortholog of Mago that was discussed briefly by Mark Johnson (Brown University). Disruption of *AtMago* results in defective pollen tube guidance. Similar to the function of Mago, *AtMago* may be important for the asymmetric localization of specific mRNAs during pollen tube growth.

John Brown (Scottish Crop Research Institute) presented exciting results from his study on the *Arabidopsis* nucleolus proteome (Pendle et al. 2005). Using tandem mass spectrometry, 217 proteins were identified from nucleoli isolated from *Arabidopsis* cell cultures. Transient expression of the full-length cDNAs of 76 of these proteins, fused to GFP, confirmed nucleolar localization for 87% of them. A comparison with the proteome of the human nucleolus (Leung et al. 2006) showed that 69% of the *Arabidopsis* nucleolar proteins have putative orthologs in the human nucleolus. Moreover, the variety of functions identified suggests several new roles for the nucleolus including mRNA splicing surveillance and nonsense-mediated decay (NMD). For instance, six proteins whose animal orthologs are involved in the EJC were found in the *Arabidopsis* nucleolar proteome. EJCs mark splice junctions after mRNA splicing and are involved in mRNA export and NMD. The role of the nucleolus in RNA surveillance and/or NMD was investigated by analysis of the nucleolar...
transcriptome. Cloning and sequencing of 226 full-length cDNAs from the nucleus showed that 38% of them correspond to aberrantly spliced transcripts while the proportions of aberrantly spliced transcripts in the whole nucleus and the whole cell are just 13% and 2%, respectively. According to the rules known to lead to NMD in animals, 67% of the aberrantly spliced transcripts identified in the Arabidopsis nucleolar transcriptome would be subject to NMD. Whether aberrantly spliced transcripts are exported to the nucleus once they are detected or whether all mRNAs go through the nucleolus for their surveillance is not known. In addition, whether these aberrant transcripts are degraded in the nucleus or exported to the cytoplasm for degradation remains to be elucidated.

An important component of eukaryotic RNA degradation is the exosome, a macromolecular machine that plays a central role in the processing, surveillance, and turnover of many RNAs (for review, see Buttner et al. 2005). Investigating the plant exosome is the major research focus of Dmitry A. Belostotsky and his laboratory (State University of New York at Albany). His laboratory has identified the Arabidopsis exosome subunits by tandem affinity purification tag-assisted proteomic and bioinformatic analyses. The different core subunits are evolutionarily conserved. The most studied Arabidopsis exosome subunits are the exonucleases RRP41 and RRP4 (Chekanova et al. 2002). These two proteins were shown to interact with each other, but analysis of rrp41 and rrp4 loss-of-function mutants revealed that they have different developmental functions. The rrp41 loss-of-function mutant is female gametophytic lethal, whereas the rrp4 loss-of-function mutant is embryo lethal. In the developing seeds of flowering plants, there is a nutritive tissue called endosperm. Before arrest, the rrp4 embryo assumes partial endosperm identity and exhibits a loss of parental imprinting of known key regulatory genes. This results from both silencing of paternal alleles and activation of maternal alleles’ expression. The effect of knocking down RRP41 or RRP4 gene expression at later stages was investigated using estradiol-inducible RNA interference (RNAi). In order to identify the RNA substrates of these two exosome subunits, whole-genome tiling microarrays were employed. This analysis revealed that the transcriptome was differentially modified (but with some overlap) by the silencing of RRP4 and RRP41. Furthermore, the affected RNAs came from five categories: ribosomal RNAs, sn/snoRNA, miRNA precursors, mRNAs, and unknown noncoding RNAs. Further investigation of these will greatly help in understanding biological functions of the plant exosome.

Additional insights into the RNA degradation processes in plants were provided by Jaimie Van Norman (University of Utah, from the laboratory of Leslie E. Sieburth) and Katarzyna Kruszka (Warsaw University). The Sieburth laboratory identified TRIDENT, the Arabidopsis ortholog of DCP2, an mRNA decapping enzyme. TRIDENT (TDT) function was investigated because tdt mutants produce a disconnected vein pattern on the seed leaves called cotyledons, similar to another Arabidopsis mutant called varicose (vcs) (Deyholos et al. 2003). The protein encoded by VCS is a homolog of a DCP2-binding protein, GE-1/HELDS, which in animals is required for assembly and activity of the mRNA decapping complex (Fenger-Gron et al. 2005; Yu et al. 2005b). Thus, VCS and TDT may function together in RNA degradation. This is supported by the observation that RNAs accumulate in much greater levels in the tdt and vcs mutants compared to wild type. This and other studies suggest that the plant 5’ to 3’ decay machinery is very similar to that of animals. Furthermore, DCP2 and GE-1/HELDS were shown in humans to interact with the cytoplasmic LSM (SM-like) complex, which is involved in RNA degradation and processing. The nuclear LSM complex is a core component of U6snRNP involved in splicing. Krusza observed a decrease in the level of U6snRNA in AtLsm5 and AtLsm8 mutants, as well as defects in the splicing of some mRNAs. Moreover, a stabilization of ribosomal protein mRNA and of mRNA containing the ARE and DST (RNA decay targeting) elements was observed in the AtLsm5 mutant (but not in AtLsm8). Therefore, AtLSM5 and AtLSM8 are likely involved in RNA splicing and AtLsm5 appears to play a role in RNA degradation, but both processes only affect a subset of mRNAs.

Recent findings about plastid rRNA processing and stability were presented by Thomas Bollenbach (Cornell University, laboratory of David Stern). In plants, the four plastid rRNAs, 16S, 23S, 4.5S, and 5S, and three tRNAs emanate from a single polycistronic primary transcript that is heavily processed. The initial step of the processing pathway is endonucleolytic cleavage followed by exonucleolytic maturation of pre-rRNA 3’ ends. Two nucleus-encoded genes from Arabidopsis that mediate this step in chloroplast rRNA processing code for RNase R (RNR1), a hydrolytic 3’ to 5’ exoribonuclease (Perrin et al. 2004; Bollenbach et al. 2005), and polynucleotide phosphorylase (PNPase), a phosphorolytic 3’ to 5’ exoribonuclease (Walter et al. 2002). In the rnr1 mutant, the levels of the mature rRNAs are variably reduced. The 5S rRNA is the most affected, as it accumulates to only 2% of the wild-type level, and although the rnr1 mutant also accumulates low levels of 3’ extended 5S precursors, they appear to be unstable in the mutant background. On the other hand, the pnp mutant, which is defective in 23S 3’ processing, accumulates unique 3’ extended 23S precursors compared to wild type and rnr1, which are stable in the pnp background. These results showed the importance of exonucleolytic reactions in the processing of rRNA.
precursors into functional mature rRNAs. Studies of the mechanism leading to pre-5S instability in mrr1 led to the discovery of an EST from tomato, which appears to be derived from a plastid-encoded antisense RNA that is complementary to 5S rRNA and tRNA\textsubscript{Arg}. The \textit{Arabidopsis} homolog of this RNA, which was named AS5, appears to overaccumulate in the plastids of \textit{pnp} and \textit{mrr1} mutants. Thus, PNPase and RNR1 appear to be involved in the biogenesis of AS5, whose overaccumulation in the \textit{mrr1} mutant could explain the dramatic drop in 5S RNA accumulation.

Another aspect of RNA processing discussed at the meeting was RNA editing. Michael Hayes (Cornell University, laboratory of Maureen R. Hanson) highlighted the importance of RNA editing (C-to-U conversion) for the correct expression of many chloroplast genes. Thirty-four editing sites are known in the tobacco chloroplast genome, 33 of which lead to an amino acid substitution. In the absence of consensus motifs around all the editing sites, their recognition by the editing machinery is not yet understood. Michael Hayes used an in vitro chloroplast editing assay to test the editing efficiency of the pcb1 C214 site embedded in a synthetic RNAs. This strategy showed that the 13 nt upstream and the 10 nt downstream of C214 are sufficient for editing. The GCCGUU sequence (and its position) in –11 to –7 is especially important.

**NATURAL RNA SENSORS**

To broaden the understanding of RNA in biology, the meeting included talks that discussed the general area of RNA sensors. Natural RNA sensors such as riboswitches and thermosensors mediate gene regulation via their ability to bind small molecules or respond to environmental conditions, causing changes in their secondary and tertiary structures that serve to regulate gene expression (Winkler and Breaker 2003; Mandal and Breaker 2004; Nudler 2006). The highly ordered secondary structure of the RNA allows for folding into a tertiary structure that contains a binding pocket for ligand recognition. The small molecules are completely surrounded by the RNA, which recognizes many features of the ligand (Batey et al. 2004; Nudler 2006). The structures are stabilized by base stacking and base pairing, i.e., hydrogen bonding, hydrophobic interactions, and van der Waals interactions. The RNA folds involve tertiary interactions including loop–loop contacts and pseudoknots that aid in the formation of the binding pocket (Thore et al. 2006).

Robert Batey described the remarkable crystal structures of the guanine and the S-adenosylmethionine riboswitches that his group has recently solved (Batey et al. 2004; Montange and Batey 2006). Although these particular riboswitches are present in bacteria, they have not yet been found in plants. The guanine riboswitch was found in the \textit{5’} UTR of \textit{Bacillus subtilis} and binds directly to guanine, hypoxanthine, or xanthine. The mechanism of transcription control is similar to the classical transcription attenuation mechanism seen in the \textit{trp} operon of \textit{B. subtilis}, which binds the TRAP protein. Ligand binding inhibits the formation of an antiterminator stem–loop, thereby promoting transcription termination (Babitzke and Gollnick 2001; Batey et al. 2004).

In addition, there is clear evidence for riboswitches in plants, as was discussed briefly by Batey. \textit{Arabidopsis}, as well as a few other plant and fungi species, carry a thiamine pyrophosphate (TPP) biosynthesis gene that has an aptamer domain. In plants, the riboswitch is located in the \textit{3’} UTR instead of the \textit{5’} UTR (Sudarsan et al. 2003). Although the function of this riboswitch is unknown, it has been suggested that it affects either mRNA stability or processing. The structure of the eukaryotic TPP aptamer domain solved by Ban and coworkers revealed a ligand-recognition conformation similar to the riboswitches from bacteria (Thore et al. 2006).

Natural RNA sensors are also involved in the response to heat shock in mammalian cells. New work in this area was reported by Evgeny Nudler (New York University Medical School). Under heat shock, the transcription elongation factor eEF1A is released from previously bound states and is thought to form a complex with a noncoding heat shock RNA (HSR1). HSR1, which changes conformation dramatically in response to even a slight change in temperature, is thought to act as a thermosensor. Together, eEF1A and HSR1 govern the trimerization of heat shock factor1 (HSF1) from its monomeric form, allowing for the activation of heat-shock protein genes (Shamovsky et al. 2006).

In addition to responding to heat, changes in RNA structures were suggested to aid in the cold response. Kentaro Nakaminanmi (West Virginia University) gave the example of a cold-shock domain protein from wheat, WSCP1, which may be involved in this process. In \textit{Escherichia coli}, cold-shock proteins (CSPs) are highly accumulated in response to cold shock (Goldstein et al. 1990), and function as transcription antiterminators or translational enhancers, by resolving cold-induced RNA processing. The structure of the eukaryotic TPP aptamer domain. In plants, the riboswitch is located in the \textit{5’} UTR instead of the \textit{3’} UTR (Sudarsan et al. 2003). Although the function of this riboswitch is unknown, it has been suggested that it effects either mRNA stability or processing. The structure of the eukaryotic TPP aptamer domain solved by Ban and coworkers revealed a ligand-recognition conformation similar to the riboswitches from bacteria (Thore et al. 2006).

**RNA-plant meeting**
SMALL RNAs INFLUENCE GENOME EVOLUTION

In addition to the involvement of RNAs in rapid responses such as environmental stress responses, several examples were given during the symposium of how RNA influences the evolution of genetic sequences. Keynote speaker David Bartel (Whitehead Institute) discussed some aspects of miRNA/siRNA influence on gene evolution. Previously, it had been shown that many messages are under selective pressure to maintain 7mer sites complementary to the seed regions of animal miRNAs (Brennecke et al. 2005; Farh et al. 2005; Krek et al. 2005; Lewis et al. 2005). By this criterion, over a third of the human genes appear to be conserved miRNA targets (Brennecke et al. 2005; Farh et al. 2005; Krek et al. 2005; Lewis et al. 2005). In mammals, there is also evidence for an interesting phenomenon called selective avoidance (Farh et al. 2005). When it is developmentally or physiologically advantageous for gene expression to be high, temporally and spatially, and evolutionarily or physiologically advantageous for gene expression to be low, a miRNA gene is coordinately expressed, any 7mer within the 3′ UTR that would correspond to the miRNA seed sequence is selected against it. By contrast, when there is no selective pressure against having that site in the 3′ UTR of that gene. An avoidance of 7mer sites in mammals indicates that these short sites are not only important but are often sufficient for repression in animals (Farh et al. 2005). No similar observations have been reported in plants. This can be explained by the idea that in plants much more extensive pairing appears to be required for miRNA recognition (Jones-Rhoades et al. 2006).

Another example of how RNA influences evolution of the genetic sequence was given by Michael Hayes from Maureen Hanson’s laboratory (Cornell University), who studies RNA editing of the chloroplast genome (see the RNA processing section). The GCCGCUU element downstream from the pBE C214 editing site is conserved in species that edit, while species that have a T in the C214 position do not have this conserved element, underscoring the effect of editing on evolution.

Similar to the selective pressure for conserving a specific recognition domain or to avoiding one, genes that code for riboswitches (see previous section) have been under selective pressure to maintain a strict secondary structure of their mRNA. The aptamer domain, or ligand binding domain, of each riboswitch has a unique secondary structure that has evolved to recognize a specific ligand. For example, the three-way junction-binding domain of a guanine riboswitch is phylogenetically conserved (Mandal et al. 2003).

METHODS OF ANALYSIS AND NOVEL TOOLS

As advances in the genomics, epigenomics, and other “-omics” fields continue, the amount of data available to be processed increases dramatically. In order to keep up with these quickly growing fields, higher throughput methods of analysis are being developed to handle this emerging cornucopia of biological information.

sRNA microarrays have been used to examine siRNAs and miRNAs within plants by Pamela Green and collaborators (see below). An expansion of microarray technology is seen in the whole genome tiling arrays that are used in transcriptome analysis, studying gene expression, gene discovery, and even in identifying targets of RNA-binding proteins. Tiling arrays, described by Joe Ecker, have the advantage that the DNA oligonucleotides span the entire genome, which allows for mapping of results in chromosome order and improves the accuracy of the genome annotations (the Arabidopsis Tiling Array Transcriptome Express Tool http://signal.salk.edu/cgi-bin/atta; Yamada et al. 2003).

RNA coimmunoprecipitation chip analysis, or RIP-Chip, also utilizes microarray analysis to identify RNAs that are associated with specific RNA-binding proteins in vivo. Jack Keene (Duke University) used this method in order to characterize the RNA components of messenger RNA–protein complexes, or mRNP. RNA-binding proteins (RBPs) associate with unique sets of RNAs, and it is therefore possible to identify RNAs involved in the same biological process or response (Keene et al. 2006). Keene described experiments from his and several other groups that used RIP-Chip and related methods to identify distinct subsets of mRNAs encoding functionally related proteins expressed in eukaryotes from yeasts to plants to mammals. These and other experiments presented support the concept of post-transcriptional RNA operons that he proposed to explain how RNA-binding proteins and noncoding sRNAs regulate mRNA processing at the levels of splicing, export, stability, and translation to form coordinated gene expression networks (Keene and Tenenbaum 2002). Kenneth Watkins (University of Oregon, laboratory of Alice Barkan) described the use of the RIP-Chip technique to identify substrates for chloroplast PPR proteins and group II intron splicing factors (Schmitz-Linneweber et al. 2005). In addition to RIP-Chip, methyl-C immunoprecipitation (mCIP) is a similar technique that has been used by Ecker to analyze methylation data and attempt to find methylation patterns within the chromosomes.

Massively parallel signature sequencing (MPSS) and pyrosequencing (Brenner et al. 2000; Margulies et al. 2005) are powerful techniques to sequence and measure expression levels of hundreds of thousands of sRNAs (Lu et al. 2005; Henderson et al. 2006). Subsequently, the sequences obtained can be quantified and mapped to the relevant genome and transcriptome (Lu et al. 2005; Henderson et al. 2006). Green, Carrington, and Bartel reported using these methods to find new miRNAs and investigate the effects of different RNA-silencing mutants. For example, Green, who pioneered the use of these
techniques for studying sRNAs, utilizes these techniques to compare the sRNA profiles between rice and *Arabidopsis*. (A database for the MPSS data, including sRNA data, is available online at http://mpss.udel.edu/at for *Arabidopsis* and http://mpss.udel.edu/rice for rice.) With Agilent Technologies, Pamela Green and Blake Meyers’ laboratories developed a DNA microarray with >2000 sRNA sequences and used it to examine tissue-preferential expression patterns in *Arabidopsis*.

Some *Arabidopsis* miRNAs currently have only one predicted (or validated) target. With such a large genome, one might expect a miRNA to have multiple targets. RNAhybrid is a program designed to predict new miRNA targets in any organism. RNAhybrid optimizes the free energy of the miRNA target’s hybridization site. The putative target miRNAs are filtered through several steps in order to minimize redundancy and to assure that hits are within specific free energy and hybridization constraints. Leonardo Alves-Junior (Bielefeld University in the laboratory of Bernd Weisshaar) described results for *Arabidopsis*, where 296 potential new targets for miRNAs in 42 families were found. Functional predictions can be made based on the Gene Ontology (GO) annotations, which are now awaiting experimental validation (http://bibiserv.techfak.uni-bielefeld.de/rnahybrid/; Rehmsmeier et al. 2004).

To assist in sRNA cloning data entry, Alex Ebhardt (Simon Fraser University) created *Ebbie*, a program assisting sRNA cloning projects with up to 1000 sequences (www.bioinformatics.org/ebbie; Ebhardt et al. 2006). The cloned sequences are flanked by 5′ and 3′ cloning primers. *Ebbie* searches for the cloning primers in order to identify and analyze the cloned sequences. The program’s output includes a unique ID, orientation, length, sample source, as well as sequence for each clone. Each sequence is further analyzed using locally installed BlastN databases, one of which is a dynamic database that allows for rapid identification of frequently cloned sequences and clusters of sRNA sequences.

RNA–protein interactions involve complicated pathways and structures that are essential for many processes in almost all organisms. RIP-Chip, described above, is one of the techniques being used to gain more insight into RNA–protein interaction networks. In order to organize information about RNA-binding proteins in plants, the Plant RNA Binding Protein/Putative Orthologous Groups Database, or PlantRBP/POGS Database, was created (http://plantrbp.uoregon.edu). In addition to storing data for RBPs, this database attempts to organize all proteins from *Arabidopsis*, rice, and maize into “putative orthologous groups” (POGs), to annotate POG members with domain content, predicted intracellular location and phylogenetic data, and to display the data for POG members in a manner that facilitates comparative analyses.

**CONCLUDING REMARKS**

This symposium gave a place of honor to sRNAs. Hence, their biogenesis, biological functions in plant biology, as well as exciting hypotheses about their impact on the evolution and organization of genomes were discussed. With at least four DCL proteins, all having specific roles, plants appear to have similar yet distinct sRNA pathways than other organisms. Roles of miRNAs in plant development have been highlighted, and they likely also play roles in stress responses, and possibly in other processes. A class of sRNAs more recently discovered, the nat-siRNAs, join them in the stress–response function. They are a fascinating group due to their overlapping localization in the genome and antagonistic expression of the two genes from which they emanate. Additional layers of control in gene expression, such as RNA processing, RNA targeting, RNA decay, and riboswitches, which were also covered during the meeting, point to wider areas of RNA studies for future growth. These compelling studies not only completed the up-to-date overview of RNA biology in plants, but also testified to the mutual nourishment between the biology of plant and nonplant systems in this cutting-edge area of research.

**ACKNOWLEDGMENTS**

We thank Sarah M. Assmann, Phil Bevilacqua, Teh-hui Kao, and Hong Ma for organizing the meeting and for their guidance and comments in writing this review, and Michael Axtell for helpful comments. We also thank all of the speakers who participated in this symposium. We gratefully acknowledge support for the symposium from the National Science Foundation (MCB-0602024) and the US Department of Agriculture (2006-01127) as well as the Eberly College of Science, the Huck Institute of Life Sciences, and the Pennsylvania State University College of Agricultural Sciences. Research on plant RNA biology in the Assmann and Bevilacqua laboratories was supported by NSF MCB-0209694. C.L.H.H. was partially supported by the Pennsylvania State University Integrative Biosciences Graduate Program and a grant from the Department of Energy (DE-FG02-02ER15332).

**REFERENCES**


The blossoming of RNA biology: Novel insights from plant systems

Jérôme Bove, Carey L.H. Hord and Melissa A. Mullen

RNA 2006 12: 2035-2046 originally published online October 19, 2006
Access the most recent version at doi:10.1261/rna.303806

References
This article cites 118 articles, 56 of which can be accessed free at:
http://rnajournal.cshlp.org/content/12/12/2035.full.html#ref-list-1

License

Email Alerting Service
Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.