Riboregulators in plant development

P. Laporte, F. Merchan, B.B. Amor, S. Wirth and M. Crespi

ISV (Institut des Sciences du Végétal), CNRS (Centre National de la Recherche Scientifique), 1 Avenue de la Terrasse, 91198 Gif-sur-Yvette, France

Abstract

npcRNA (non-protein-coding RNAs) are an emerging class of regulators, so-called riboregulators, and include a large diversity of small RNAs [miRNAs (microRNAs)/siRNAs (small interfering RNAs)] that are involved in various developmental processes in plants and animals. In addition, several other npcRNAs encompassing various transcript sizes (up to several kilobases) have been identified using different genomic approaches. Much less is known about the mechanism of action of these other classes of riboregulators also present in the cell. The organogenesis of nitrogen-fixing nodules in legume plants is initiated in specific root cortical cells that express the npcRNA MTENOD40 (Medicago truncatula early nodulin 40). We have identified a novel RBP (RNA-binding protein), MtRBP1 (M. truncatula RBP 1), which interacts with the MTENOD40 RNA, and is exported into the cytoplasm during legume nodule development in the region expressing MTENOD40. A direct involvement of the MTENOD40 RNA in the relocalization of this RBP into cytoplasmic granules could be demonstrated, revealing a new RNA function in the cell. To extend these results, we searched for npcRNAs in the model plant Arabidopsis thaliana whose genome is completely known. We have identified 86 novel npcRNAs from which 27 corresponded to antisense RNAs of known coding regions. Using a dedicated 'macroarray' containing these npcRNAs and a collection of RBPs, we characterized their regulation in different tissues and plants subjected to environmental stress. Most of the npcRNAs showed high variations in gene expression in contrast with the RBP genes. Recent large-scale analysis of the sRNA component of the transcriptome revealed an enormous diversity of siRNAs/miRNAs in the Arabidopsis genome. Bioinformatic analysis revealed that 34 large npcRNAs are precursors of siRNAs/miRNAs. npcRNAs, which are a sensitive component of the transcriptome, may reveal novel riboregulatory mechanisms involved in post-transcriptional control of differentiation or environmental responses.

Introduction

In eukaryotic cells, the complex responses to external factors are governed by post-transcriptional processes allowing rapid cellular adaptations to the new conditions. RNAs can exert structural functions (such as tRNA and the rRNA) as well as regulatory roles. npcRNAs (non-protein-coding RNAs) are generally a class of RNAs that do not encode proteins but instead may function as RNA molecules. As protein-coding genes, npcRNAs are a heterogeneous group and have been divided into different classes according to their length and/or function. With respect to length, npcRNAs can range from 18 to 25 nt for the families of miRNAs (microRNAs) and siRNAs (small interfering RNAs), 20–300 nt for sRNAs commonly observed as transcriptional and translational regulators in different RNP (ribonucleoprotein) complexes (e.g. snoRNAs [small nucleolar RNAs]) and signal recognition particle or SRP RNA, up to and beyond 10000 nt for medium and large RNAs involved in other processes, including localized translation, silencing or gene inactivation [1]. We used the term npcRNAs instead of non-coding RNAs, which seems more appropriate as every sequence is coding and certain of these npcRNAs may encode small oligopeptides encoded in sORFs (short open reading frames). These sORFs should not be neglected since they could be translated into specific conditions as has been shown for the rRNA, a canonical RNA, in Escherichia coli [2]. Indeed, several RNAs originally described as non-coding were shown to code for small peptides and vice versa [3].

npcRNA genes constitute a so far underexplored component of the transcriptome, because due to the small sizes of the encoded sORFs, they have eluded bioinformatic searches. In recent years, a great number of novel npcRNA candidates have been identified in various model organisms from E. coli to Homo sapiens using both experimental and computational screenings (e.g. [4,5]). Results from the recently developed tilling arrays and systematic sequencing of full-length cDNA libraries indicated that much larger portions of eukaryote transcriptomes represent non-protein-coding transcripts than previously believed [6–8]. The recent advent of genomic approaches in model organisms has revealed a large diversity of npcRNAs, including a surprising number of antisense
RNA transcripts, pseudogenes and truncated transcripts [9]. Interestingly, several of these npcRNAs are induced at specific stages of development or during responses to environmental stresses in diverse multicellular eukaryotes [7,10]. More generally, npcRNAs have been shown to play roles in a variety of post-transcriptional processes such as the regulation of mRNA translation, stability, modification or localization, particularly targeting mRNAs coding for regulators of a wide variety of processes (e.g. transcription factors [11,12]). However, classical genetic approaches revealed few npcRNAs linked to development. This led to the proposal that riboregulators may fine-tune mRNA levels in the cell and play a more critical role in the canalization of developmental processes during organism growth rather than in differentiation itself [13].

An emerging class of post-transcriptional regulators are small npcRNAs

One of the most studied npcRNA species are single-stranded 20–27 nt sRNAs (small RNAs) belonging to two classes, miRNAs and siRNAs, known to play essential roles in the four Eukaryote kingdoms (ciliates, fungi, plants and animals). miRNAs are 20–24-nt-long RNAs, initially discovered in Caenorhabditis elegans as temporal regulators of larval differentiation and more recently in mammals and plants [11,14–18]. miRNAs are encoded by particular genes generally present, in plants, in intergenic regions of the genome. Maturation of the primary transcript, generated by RNA polymerase II, requires the intervention of a particular type III RNase named DCL1 (DICER-like 1), which cleaves twice on a hairpin-structured dsRNA (double-stranded RNA) [19]. The mature miRNA is then incorporated into a protein complex, the so-called RISC (RNA-induced silencing complex) that can recognize mRNAs partially or completely complementary to the miRNA nucleotide sequence. This recognition event mediated by the RISC-loaded miRNA leads to the cleavage (as is generally the case in plants) or the translational inhibition of the target mRNA (as is generally the case in animals). Up to now, 43 miRNA families in 71 different plant species have been defined using homology criteria [20]. Sequences of certain miRNAs families as well as their targets are highly conserved, suggesting that those miRNAs may play the same function in different species. Nevertheless, many other miRNAs are specific to only one or few phylogenetically related species, indicating their rapid evolution. In plants, miRNAs have been shown to play significant roles notably in the regulation of differentiation and in response to environmental conditions [18]. miRNAs action is exerted directly on transcripts coding for genes involved in development (e.g. the ARF (ADP-ribosylation factor) transcription factors [21] and plant stress responses [22–25]). Indeed, several miRNAs were shown to be regulated by abiotic stresses (cold, drought and salt stresses) or ABA (abscisic acid) treatments [26].

Another class of sRNAs is the siRNAs, initially identified in plants [27]. They intervene mainly in two processes: changes in chromatin conformation (e.g. through methylation) and destruction of foreign RNAs such as viral RNAs or aberrant transgene mRNAs [28]. Plant siRNAs are 21–24 nt RNAs generated from long perfectly matched dsRNAs by the action of DCL2 and DCL3 enzymes [29]. These siRNAs lead to the silencing, either post-transcriptionally [PTGS (post-transcriptional gene silencing)] or transcriptionally [TGS (transcriptional gene silencing)], of expression of the gene from which the dsRNA originates. In plants, two other endogenous pathways leading to gene silencing have been described, one mediated by the tasiRNAs (trans-acting siRNAs) and one mediated by nat-siRNAs (natural antisense-mediated siRNAs) [18,30]. The 21 nt-long tasiRNAs are different from the siRNAs due to their action in trans on a gene different from the one encoding them [31,32]. A long npcRNA is generated from a TAS locus, which is itself cleaved by the action of a miRNA on one or two sites [33]. The npcRNA cleavage products are recognized by an RNA-dependent RNA polymerase and form a dsRNA, which becomes a substrate of DCL4 producing the 21 nt tasiRNAs. The identification of nat-siRNAs revealed a new mechanism involved in Arabidopsis stress responses [30]. Under salt stress conditions, a 24 nt-long siRNA could be detected, coming from two partially overlapping mRNAs that are in antisense configuration. A dsRNA formed by complementarity between a constitutively expressed gene (P5CDH) and an antisense stress-inducible transcript, is processed into these so-called nat-siRNAs. Together with secondary 21 nt siRNAs, they will induce the cleavage of P5CDH transcripts, acting thus as true siRNAs, and lead to the accumulation of an osmolyte involved in stress responses. In Arabidopsis, where the estimates of overlapping genes (potential sense/antisense RNAs) are approx. 2000, such a nat-siRNA-mediated regulation could have a strong impact on a variety of conditions. Due to the large diversity of these novel small regulatory RNAs (Figure 1), we are only beginning to understand the wide variety of processes that may be post-transcriptionally controlled by them [34,35].

A nuclear RNP is relocalized by the npcRNA MtENOD40 Medicago truncatula early nodulin 40) into the cytoplasm of Arabidopsis protoplasts

Cellular RNAs associate with RBPs (RNA-binding proteins) to form RNP particles involved in processing, nucleocytoplasmic transport, localization, translation and/or stability of mRNAs. The protein components of the RNPs (or RBPs) are thought to play crucial roles in those processes [36,37]. The RNA molecules forming part of these RNP complexes may have novel intracellular roles as shown for the intron-encoded snoRNAs [38], the miRNAs [11] or the BC1 transcript in mammals [39]. The RNA molecule may determine the functional specificity of the complex and analysis of RNA–protein interactions may reveal novel RNA functions in these cellular complexes. npcRNAs may
Diverse cellular RNAs may be involved in post-transcriptional regulation of gene expression

npcRNAs are encoded in the genome and may produce a myriad of sRNAs (nat-siRNAs, miRNAs and tasiRNAs). The npcRNAs or their derivatives may then regulate mRNA stability or translation of specific mRNAs. This regulation is mediated by RBPs and modulates the different cellular functions required for plant growth and development.

constitute examples that have retained specific RNA functions linked to their RNA sequences or structures independently of their translational capabilities [12].

The npcRNA MtENOD40 is expressed at the onset of root nodule organogenesis, a legume-specific process. Its expression pattern follows the differentiation gradient from the initial dividing cells (or meristematic cells) to the fully differentiated cells of the nodule. The MtENOD40 RNA is highly structured [40,41] and a role of the RNA moiety in gene function has been proposed [42]. To investigate cellular functions of the MtENOD40 RNA, potential protein partners of the transcripts have been sought using the yeast three-hybrid system. We have identified a novel nuclear RBP, MtRBP1 (M. truncatula RBP 1), which localizes in nuclear speckles [43]. Immunolocalization experiments and transient assays demonstrated that relocalization of the nuclear MtRBP1 into the cytoplasm is a novel function of the MtENOD40 npcRNA during nodule organogenesis in legumes. This relocalization event could be induced in Arabidopsis protoplasts, a heterologous system (Figure 2). Introduction of the MtENOD40 RNA relocalized MtRBP1 into cytoplasmic particles, revealing that this mechanism is conserved even in non-legume plants.

npcRNAs are a sensitive component of the Arabidopsis transcriptome

These results prompted us to search for npcRNAs in the model plant Arabidopsis. We have identified 43 npcRNAs in A. thaliana [10] and we extended this approach to a recently characterized full-length cDNA library [44]. A total of 86 novel npcRNAs of which 27 corresponded to antisense RNAs of coding genes were identified. The expression of these genes was studied using a ‘dedicated’ macroarray (RIBOCHIP) containing these npcRNAs and a collection of genes encoding a variety of RBPs belonging to all major groups from Arabidopsis. We hybridized the RIBOCHIP with probes obtained from total RNA from different tissues of A. thaliana Col0 grown under a variety of conditions including Arabidopsis leaves compared with roots, inflorescences of dcl1-9 mutant compared with wild-type and plant roots grown under different abiotic stresses. Of the 42 genes detected as differentially expressed in at least one of the conditions assayed, 27 correspond to npcRNAs. The observed bias towards the npcRNAs (65% of the regulated genes) points to a dynamic regulation of these transcripts over RNA-related protein-coding genes, whose regulation can be additionally controlled at translational and post-translational level. Several npcRNAs were regulated in more than one condition, suggesting a role in diverse processes. Since a large number of sRNAs have been recently identified ([35,45]; http://asrp.cgrb.oregonstate.edu/db/), we mapped our 86 candidate npcRNAs in databases and found that 34 of them present a hit with at least one siRNA up to more than a hundred of sense or antisense siRNAs. One example of such npcRNA, the npc83, is shown in Figure 3. Hence, several npcRNAs turned out to be precursors of sRNAs and gave us insight into the complexity of siRNA processing in plants.

Concluding remarks

In addition to the diverse mechanisms implied in growth control, post-transcriptional regulation of developmental regulators (such as transcription factors) mediated by npcRNAs (such as tasiRNAs/miRNAs) is emerging as an important determinant of differentiation in eukaryotes. These riboregulatory mechanisms may be particularly relevant...
Figure 3 | Example of an npcRNA containing several sRNAs in a highly structured region

The npc83 [10] encodes several siRNAs/miRNAs detected in ASRP databases (http://asrp.cgrb.oregonstate.edu/db/). The predicted secondary RNA structure of npc83 is shown and the dsRNA region involved in the production of several siRNAs/miRNAs is magnified together with the sequences detected in databases.

to adjusting differentiation processes to the environmental conditions encountered during growth and hence may play significant roles in plant developmental plasticity.

P.L. is supported by the MENR (Ministère de l‘Education Nationale, de l’Enseignement Supérieur et de la Recherche). F.M. is supported by the Ministerio de Educacion y Ciencia. The support from RIBOREG FP6 EU (European Union) project is also acknowledged.

References

©The Authors Journal compilation ©2007 Biochemical Society

Received 26 June 2007
doi:10.1042/BST0351638