The Diversity of Plant Small RNAs Silencing Mechanisms

Jens A. Schröder and Pauline E. Jullien*

Abstract: Small RNAs gene regulation was first discovered about 20 years ago. It represents a conserve gene regulation mechanism across eukaryotes and is associated to key regulatory processes. In plants, small RNAs tightly regulate development, but also maintain genome stability and protect the plant against pathogens. Small RNA gene regulation in plants can be divided in two canonical pathways: Post-transcriptional Gene Silencing (PTGS) that results in transcript degradation and/or translational inhibition or Transcriptional Gene Silencing (TGS) that results in DNA methylation. In this review, we will focus on the model plant Arabidopsis thaliana. We will provide a brief overview of the molecular mechanisms involved in canonical small RNA pathways as well as introducing more atypical pathways recently discovered.

Keywords: Argonaute · DNA methylation · Plant · Silencing · Small RNA

1. Introduction

Eukaryotes have evolved very complex mechanisms to regulate gene expression during organism development as well as to respond to stresses. One of these mechanisms is based on small non-coding RNAs (sRNA). In plants, sRNAs regulate gene expression by transcriptional inhibition referred as Transcriptional Gene Silencing (TGS) or by Post-transcriptional Gene Silencing (PTGS) where sRNAs lead to downregulation of messenger RNA (mRNA) level or inhibition of protein translation.

First evidence for TGS and PTGS in plants arose in the early nineties when plant scientists started to engineer genetically modified organisms (GMO). At the time, GMOs were made by introducing a custom-made Transfer DNA (T-DNA) into the plant genome using Agrobacterium-mediated plant transformation. In introducing a custom-made T-DNA, the introduction of the second T-DNA led to the ‘inactivation’ of the first T-DNA despite being in a different chromosomal location (i.e. trans-silencing). They further demonstrated that this ‘inactivation’ of the first T-DNA was correlated with DNA methylation level. This phenomenon was later shown to rely on a, nowadays, well-characterised pathway called RNA-directed DNA methylation (RdDM).[2] RdDM is the main small RNA pathway involved in TGS in plants. Similarly, in 1990, scientists trying to overexpress a petunia pigmentation gene (CHS) to obtain deep purple flowers observed unexpected colours ranging from intense purple to marble to fully white flowers. By overexpressing CHS, they observed silencing of both the T-DNA encoded CHS and the endogenous CHS. The authors referred to this phenomenon as ‘co-suppression’. A few years later, in 1993, tobacco plants over-expressing a viral coat protein from tobacco etch virus (TEV) were found to be resistant to viral infection.[5] This resistance was limited to TEV and closely related viruses and therefore relied on sequence homology. Molecular analysis revealed that silencing was due to low transcript level despite a stable transcription rate which defines PTGS. The involvement of sRNA in these observations came later when RNA interference was discovered in nematodes[6] and sRNA were found to be associated with several silenced GMO plants.[7]

The basic mechanism of RNA interference or sRNA-mediated gene regulation is highly conserved from plants to animals. It relies on an initiator, the sRNA, which is cleaved from double-stranded RNA by an RNA endonuclease called DICER. In the model plant Arabidopsis thaliana, four genes encode for DICER-LIKE proteins (DCL1 to 4) and they generate sRNAs of different lengths (21-nt, 22-nt or 24-nt).[8] After dicing, the resulting sRNA is loaded into an AGONAUTE (AGO) protein, the effector of the sRNA pathways. Arabidopsis encodes 10 AGOs which differs in specificity and substrate preferences and is loaded into an ARGONAUTE (AGO) protein, the effector of the sRNA pathways. Arabidopsis encodes 10 AGOs which differs in specificity and substrate preferences.

2. Post-transcriptional Gene Silencing (PTGS) Pathways

As previously mentioned, PTGS is achieved either by mRNA cleavage or translational repression. Initiator sRNA can be divided
into two major groups, microRNA (miRNA) and small interfering RNA (siRNA). We will describe here, first, the miRNA pathway and then two specific classes of siRNA: tasiRNA and vsiRNA.

### 2.1 microRNAs

In plants and other organisms, miRNA are essential for proper development\(^{19}\). As a consequence, mutants affected in the miRNA pathway display strong developmental phenotypes. Mutations affecting the main actors of the miRNA pathway such as DICER (DCL1) and AGO1, show phenotypes ranging from sterility, embryo lethality, flower defects to leaf development phenotypes depending on the impact of the mutation on protein function.\(^{11,12}\)

For example, AGO1 and its closest homologue AGO10 play an important role in shoot apical meristem regulation. Plant shoot apical meristem represents a stem cell niche from which all plant organs are generated. Plant stem cells require the expression of HD-ZIP III transcription factors which are regulated by miRNA miR166/165. In non-stem cells, AGO1 loaded with miR166/165 represses HD-ZIP III. However, in stem cells, AGO10 preferentially binds to miR166/165 preventing AGO1 mediated repression of HD-ZIP III.\(^{13,14}\) In addition to their developmental role, some miRNAs have been shown to also be involved in plant resistance to stresses.\(^{15}\)

The first step of miRNA biosynthesis is the transcription of a primary miRNA transcript (pri-miRNA) by RNA Polymerase II (Pol II). The pri-miRNA imperfectly self-folds into a hairpin structure which is further processed in two steps: (i) from pri-to-precursor (pre)-miRNA and, (ii) from pre-to-mature miRNA, by DCL1.\(^{16}\) Dicing takes place in the nucleus, where DCL1 can be localized in the so-called dicing body.\(^{17}\) The cleavage of the microRNA hairpin by DCL1 is assisted by additional proteins such as Hypenastis leaves 1 (HYL1) or Serrate (SE).\(^{18}\) The double-stranded diced product, usually of 21 nucleotides in length, is methylated at the 3’ end on both strands by HUA ENHANCER 1 (HEN1). Methylation of the miRNA at the 3’ ends prevents its degradation.\(^{19}\) The double-stranded mature miRNA is loaded into AGO1. The AGO1 loaded strand is called the guide (complementary to the target RNA) strand while the released strand is called the passenger or star strand (miR*).\(^{18}\) miR* are generally unstable and degraded except in some cases where the miR* can be loaded in a different AGO protein such as miRNA393b*.\(^{20}\) In contrast to animal systems, in plants, AGO1 has been recently shown to shuttle between the cytoplasm and the nucleus.\(^{21}\) In the current model, unloaded AGO1 is targeted to the nucleus where it will load the mature miRNA. Loaded AGO1 will then shuttle back to the cytoplasm to perform PTGS (Fig. 1). As previously mentioned, PTGS is achieved either by mRNA cleavage or translational repression. It is yet not fully understood how these two mechanisms are regulated. For most Arabidopsis AGOs, their cleavage ability is required for their function\(^{22,23}\) leading to the conclusion that cleavage might be the preponderant mode of action of plant AGOs. However, despite being more challenging to demonstrate, several studies have shown AGO role in translational repression.\(^{24,25}\) The distinction between both mechanisms of repression could be linked to the strength of the pairing existing between the miRNA and its target, AGO interacting partners, as well as potential differences in sub-cellular localisation.

### 2.2 Trans-acting siRNA

Trans-acting RNAs (tasiRNA) are peculiar small interfering RNAs found in plants. They are the product of a miRNA directed cleavage of a long non-coding RNA transcript called TAS. TAS long non-coding RNAs are encoded in the plant genome and transcribed by RNA Pol II. tasiRNA are produced by an initial miRNA targeted cleavage of a TAS transcript. The initial miRNA cleavage relies on AGO1 and/or AGO7. The double-stranded diced product, usually of 21 nucleotides in length, is methylated at the 3’ end on both strands by HUA ENHANCER 1 (HEN1). Methylation of the miRNA at the 3’ ends prevents its degradation.\(^{19}\) The double-stranded mature miRNA is loaded into AGO1. The AGO1 loaded strand is called the guide (complementary to the target RNA) strand while the released strand is called the passenger or star strand (miR*).\(^{18}\) miR* are generally unstable and degraded except in some cases where the miR* can be loaded in a different AGO protein such as miRNA393b*.\(^{20}\) In contrast to animal systems, in plants, AGO1 has been recently shown to shuttle between the cytoplasm and the nucleus.\(^{21}\) In the current model, unloaded AGO1 is targeted to the nucleus where it will load the mature miRNA. Loaded AGO1 will then shuttle back to the cytoplasm to perform PTGS (Fig. 1). As previously mentioned, PTGS is achieved either by mRNA cleavage or translational repression. It is yet not fully understood how these two mechanisms are regulated. For most Arabidopsis AGOs, their cleavage ability is required for their function\(^{22,23}\) leading to the conclusion that cleavage might be the preponderant mode of action of plant AGOs. However, despite being more challenging to demonstrate, several studies have shown AGO role in translational repression.\(^{24,25}\) The distinction between both mechanisms of repression could be linked to the strength of the pairing existing between the miRNA and its target, AGO interacting partners, as well as potential differences in sub-cellular localisation.

---

**Fig. 1.** Schematic description of the Arabidopsis miRNA biosynthesis pathway. Pol II transcribes a hairpin precursor which is diced in two steps to mature-miRNA. Mature-miRNAs become methylated by HEN1 and are loaded into AGO1. The guide strand remains in a complex with AGO1, while the passenger strand is degraded. The miRNA/AGO1 complex is exported from the nucleus (grey) into cytoplasm (white) to perform either translational repression at the endoplasmic reticulum (ER) or cleavage leading to mRNA degradation. abbreviations: Pol II, RNA polymerase II; DCL1, DICER-LIKE 1; HEN1, HUA ENHANCER 1; AGO1, ARGONAUTE 1; ER, Endoplasmic reticulum.
script is processed by one of the plant RNA-dependent RNA polymerases (RdRP), RDR6, into a double-stranded RNA (dsRNA) precursor which will then be processed by DCL4 into 21-nt tasiRNA. The resulting tasiRNA are, as miRNA, methylated by HEN1 and loaded into AGO1. In contrast to miRNA, as tasiRNA. The resulting tasiRNA are, as miRNA, methylated by RdDM. The recruitment of RDR6, as well as SGS3 and AGO7, leads to double-stranded RNA, which can be again diced by DCL3 and methylated by HEN1. In phase II, mature sRNAs are loaded into AGO1, AGO3, AGO4, AGO5, AGO7 and AGO10 have only a minor effect. Additionally, not only RDR6 is involved in dsRNA production in response to a TuMV infection, but also RDR1 and RDR2. Similar diverse usage of the DICER also occurs during viral infection. These differences/preferences could be due to different expression levels or patterns of the above-mentioned AGO/DICER/DCL proteins upon infection.

3. Transcriptional Gene Silencing (TGS) Pathways

In Arabidopsis, sRNA-induced TGS relies principally on DNA methylation at cytosine residues by the RNA-directed DNA methylation pathway (RdDM). The RdDM is key to perform de novo DNA methylation whereas other DNA methyltransferases such as MET1 or CMT3 are involved in DNA methylation maintenance during cell division. As a consequence, mutants affecting the RdDM pathway only display very mild phenotypes in laboratory control growth conditions. The role of the RdDM pathway is made very clear when there is a requirement to methylate a naïve DNA sequence such as T-DNAs, transposable elements (TE) or even DNA viruses. The RdDM pathway is therefore important in the protection of genome stability during the plant life and through generations.

The molecular mechanism involved in canonical RdDM can be divided into two phases (Fig. 2). The

\begin{figure}
\centering
\includegraphics[width=\textwidth]{fig2}
\caption{Fig. 2. Representation of the canonical RNA-directed DNA methylation pathway. Phase I representing the sRNA generation is shaded in grey and Phase II representing the DNA methylation step is in white. During phase I, the complex of SSH and CLSY1 recognises DNA methylation and recruits Pol IV. Pol IV and RDR2 produce 25-45-nt dsRNAs which are diced by DCL3 and methylated by HEN1. In phase II, mature siRNAs are loaded into either AGO4 or AGO6 and bind to the scaffold RNA produced by Pol V with the help of SUV2/9. DRM2 recruitment leads to DNA methylation. Abbreviations: SSH, SAWADEE HOMEODOMAIN HOMOLOGUE 1; CLSY1, CLASSY1; DDR, chromatin remodelling complex; Pol IV, RNA polymerase IV; RDR2, RNA-dependent RNA polymerase2; DCL3, DICER-LIKE3; HEN1, HUA ENHANCER1; AGO, ARGONAUTE; DRM2, DOMAINS REARRANGED METHYLTRANSFERASE1; dsRNA, double-stranded RNA.}
\end{figure}
first phase provides the initiator of RdDM, i.e. 24-nt siRNAs and the second is the addition of cytosine methylation.

During the first phase, the RNA polymerase IV (Pol IV) is recruited to the targeted loci. Pol IV recruitment is enhanced by the SAWADEE HOMEODOMAIN HOMOLOGUE 1 (SSH1) in interaction with CLASSY (CLSY1). Pol IV transcribes a 25-45 nt long precursor RNA that will be processed by RDR2 into a dsRNA. This precursor dsRNA will then be diced by DCL3 into 24-nt siRNA. Similar to miRNA and tasiRNA, these 24-nt siRNAs are stabilised by the methylation of their 3’ end by HEN1. Mutations affecting Pol IV, RDR2 and DCL3 therefore result in a strong decrease in 24-nt siRNAs and a decrease of DNA methylation at the targeted loci. During phase two, the locus is transcribed by a second RNA polymerase, RNA Polymerase V (Pol V). Pol V recruitment and activity are not fully understood, but some contributing proteins have been identified. SUVH2 and SUVH9 can bind to methylated DNA and recruit Pol V. In addition, the chromatin remodelling complex (DDR) is also involved and consists of DEFECTIVE IN MERISTEM SILENCING 3 (DMS3), RNA-DIRECTED DNA METHYLATION 1 (DRM1) and DEFECTIVE IN RNA-DIRECTED DNA METHYLATION 1 (DRD1). This complex was shown to be important for transcription of the scaffold RNA at specific loci. The siRNA provided by the first phase of the RdDM is incorporated into the main effector of the RdDM pathway AGO4. This loading most likely occurs in the cytoplasm and then the siRNA/AGO4 complex is relocated to the nucleus. The siRNA/AGO4 complex targets Pol V scaffold RNA in a sequence specific manner, leading to the recruitment of the de novo DNA methyltransferase DRM2 by a possible direct protein interaction between DRM2 and AGO4. Recruitment of DRM2 causes local cytosine methylation. Mutations in AGO4 show a global decrease in DNA methylation. However, this decreased methylation does not affect certain loci, most likely due to a partial redundancy between AGO4 and AGO6. Indeed, RdDM loci in Arabidopsis seem to rely on both AGO4 and AGO6 independently as well as redundantly on some loci, which is not yet fully understood.

4. Alternative and New Pathways

In the recent years, more and more instances of alternative pathways to the canonical PTGS and TGS were found. The best understood ones are actually some cases of hybrid PTGS and TGS pathway where we can find alternative sources of ‘initiator’ and functions of the ‘effector’. New pathways were also identified such as one involving DNA damage response and transcriptional activation.

4.1 ‘Hybrid’ Pathways

An alternative source of the ‘initiator’ sRNA occurs when the sRNA biogenesis pathway does not occur in a canonical manner. For example, transcription by the RNA Pol II could lead to 24-nt production by DCL3 and subsequent DNA methylation. In this case, the dsRNA template comes from Pol II transcription rather than Pol IV transcription and reverse transcription by RDR2 found in classical RdDM. Additionally, some 24-nt miRNA, transcribed by Pol II, were found to rely on DCL3 instead of DCL1 for their biogenesis. Their loading in AGO4 results in DNA methylation. More interestingly some miRNA seem to be able to give rise to either 21-nt (DCL1 cleavage) or 24-nt (DCL3 cleavage) resulting then in either PTGS or TGS. As a consequence, the processing of the precursor of such miRNA should be regulated precisely to switch from PTGS to TGS regulation of the target. Another alternative to canonical RdDM was also found to be involved in Transposon (TE) silencing in which the TE transcripts were converted into dsRNA by RDR6 and cleaved by DCL3 to result in 24-nt mediated DNA methylation. At the end of this spectrum of alternative initiators, there have been a few publications describing DICER independent RdDM where the source of the 24-nt sRNA is thought to depend on RNA exonuclease rather than DCL RNA endonuclease.

In hybrid pathways involving alternative effectors, some AGOs adopt an alternative mode of action depending on their cargo size or interacting partners. One example of such an effector alternative pathway is referred as the NERD-DNA methylation pathway. NERD stands for ‘Needed for RDR2-independent DNA methylation’. It relies on the production of dsRNA by RDR1 and RDR6. DNA methylation on the targeted loci relies on the NERD protein and AGO2/AGO1 instead of the canonical RdDM effectors AGO4/AGO6, most likely involving 21-nt sRNAs and not the 24-nt sRNAs, signature of RdDM. It is worth mentioning that three out of the ten AGOs of Arabidopsis show more relaxed binding affinity such as AGO5 which has been shown to bind different sRNA sizes (Table 1). Such AGOs can potentially play important roles in bridging different sRNA pathways. For example, AGO6 loading of 21-22-nt sRNA was involved in the DNA methylation of transposable elements fragments. The biogenesis of those sRNAs was dependent on RDR6 and DCL2/4. As mentioned previously, Arabidopsis encodes 10 AGO proteins, each of which displays different loading affinity for sRNA in terms of length and 5’ nucleotide (Table1). Such differences in sRNA loading

<table>
<thead>
<tr>
<th>Clade</th>
<th>21nt</th>
<th>22nt</th>
<th>24nt</th>
<th>3’nt</th>
<th>sRNA enriched identity</th>
<th>Mode of action</th>
<th>Refs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clade I</td>
<td>AGO1</td>
<td>+++</td>
<td>-</td>
<td>U</td>
<td>miRNA</td>
<td>PTGS</td>
<td>[63]</td>
</tr>
<tr>
<td></td>
<td>AGO10</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>A</td>
<td>miR165/miR166</td>
<td>miRNA degradation</td>
</tr>
<tr>
<td></td>
<td>AGO5</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>C</td>
<td>intergenic</td>
<td>?</td>
</tr>
<tr>
<td>Clade II</td>
<td>AGO2</td>
<td>+++</td>
<td>-</td>
<td>-</td>
<td>A</td>
<td>tasiRNA</td>
<td>PTGS/TGS</td>
</tr>
<tr>
<td></td>
<td>AGO3</td>
<td>+</td>
<td>-</td>
<td>+++</td>
<td>A</td>
<td>intergenic and TE</td>
<td>PTGS/TGS</td>
</tr>
<tr>
<td></td>
<td>AGO7</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>A</td>
<td>miR390</td>
<td>PTGS</td>
</tr>
<tr>
<td>Clade III</td>
<td>AGO4</td>
<td>-</td>
<td>-</td>
<td>+++</td>
<td>A</td>
<td>Repeats and heterochromatin</td>
<td>TGS</td>
</tr>
<tr>
<td></td>
<td>AGO6</td>
<td>+</td>
<td>++</td>
<td>A</td>
<td>Repeats and heterochromatin</td>
<td>TGS</td>
<td>[62,67]</td>
</tr>
<tr>
<td></td>
<td>AGO8</td>
<td>-</td>
<td>-</td>
<td>na</td>
<td>na</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AGO9</td>
<td>-</td>
<td>-</td>
<td>+++</td>
<td>A</td>
<td>Repeats and heterochromatin</td>
<td>TGS</td>
</tr>
</tbody>
</table>
preferences should lead to a very complex stochiometric competition and/or exclusivity between the different AGOs and siRNA sequences allowing the precise tuning of plant sRNA pathways.

4.2 sRNA-mediated DNA Repair

Aside sRNA’s involvement in PTGS and TGS, there is growing evidence in plant and animal systems that sRNAs are involved in the repair of double-strand breaks (DSB). In order to investigate the role of sRNA pathways in DSB repair, Wei et al. used a transgenic reporter system inducing DSB by expressing a DNA endonuclease. DSBs lead to the accumulation of 21-nt sRNA around the DSB lesion site named DSB-induced sRNAs (diRNA). The production of diRNA requires the coordinated action of DCL2, DCL3 and DCL4, the RNA polymerases Pol IV as well as RDR2 and RDR6. All of these actors are necessary for proper DSB repair. In their experimental set-up, the authors showed that the diRNA are loaded in AGO2, a necessary AGO for DSB repair. This study provided the first evidence that sRNAs are involved in DSB DNA repair. AGO2 was also shown to be required for DNA repair after gamma ray-induced DSBs. More recently, repair upon UV-induced DNA damage was also reported to be guided by sRNA of 21-nt in size. A genome-wide analysis of UV-induced DNA lesions revealed a strong enrichment of 21-nt sRNA to the damaged sites. These sRNAs are referred as UV-C induced small RNA (uviRNA). Schalk et al. propose a model where Pol IV transcribes ssRNA from UV-C damaged loci. The transcript is then converted into dsRNA by RDR2/RDR6 and cleaved by DCL4 into 21-nt uviRNA. The 21-nt uviRNA are then loaded into AGO1. Following UV-C irradiation, AGO1 can interact with the DNA DAMAGE-BINDING PROTEIN 2 (DDB2) and most importantly, both are present on the chromatin at the lesion sites. The authors propose that this dynamic interaction of the AGO1/DDB2 complex to the DNA UV damaged loci results in the activation of DNA repair. Although it is now clear that sRNAs are involved in DSB repair, the complexes and pathways involved seem to vary depending on the type of DNA damage.

4.3 sRNA-activated Transcription

Although most of the roles of small RNAs in plants are associated with gene silencing, one recent study has uncovered an unexpected function of Arabidopsis AGO1. AGO1 was shown to activate transcription upon hormonal inductions and stresses. Despite its main localisation and place of action in the cytoplasm, AGO1 was found to be associated to the chromatin of around 940 loci. The guiding of AGO1 to chromatin relies on 21-nt sRNA as the association of AGO1 to the targeted loci is decreased in the dcl1 mutant. A positive regulation of chromatin-bound AGO1 on transcription was observed when comparing the transcript abundance between the control and ago1 mutants. In addition, investigation of active Pol II abundance on AGO1 bound loci showed that AGO1 has a direct effect on transcription and suggests that AGO1 might facilitate the recruitment of Pol II to the locus. How AGO1 acts to regulate transcription is not fully understood, but the direct interaction of AGO1 with a SWI/SNF chromatin remodelling complex could facilitate Pol II accessibility to chromatin. Interestingly, Liu et al. further showed that association of AGO1 to chromatin as well as associated transcriptional activation is regulated by plant hormones, abiotic and biotic stimuli.

5. Conclusion

The sRNA pathways play a major role in plant development such as organ polarity but also plant pathogen interaction and responses to abiotic stresses like salt or UV radiation. As we gain more and more molecular understanding of the canonical pathways of PTGS and TGS, additional interplays between these pathways are being discovered. Some of these additional pathways are hybrids between known pathways while others are more unusual. As mentioned, the biogenesis of sRNA can be from different sources and the AGO effectors have different loading affinities and modes of action. Other layers of complexity can be added if we incorporate the sub-cellular localization of the proteins involved but also their expression level and pattern during plant development. For example, small RNA pathways unique to sexual reproduction seem to involve non-ubiquitously expressed proteins such as AGO9/2/3, DRM1 and a nuclear DCL4 isoform. In the near future, efforts should be focused on getting more insights into the cell specificity of sRNAs and their associated pathways as single-cell and cell-specific ‘omics’ experimental approaches are becoming more accessible. The intricacy of plant sRNA function and regulation also relies in their mobility. Indeed, unlike in animal models, plant sRNAs are well known to move from cell-to-cell as well as long distances through the plant vasculature which provides yet another potential layer of regulation and could potentially connect stresses affecting the mother plant with gene regulation of the next generation, a feature that deserves future investigation given the rapidly changing environmental conditions experienced worldwide.

Acknowledgements

This work was supported by a grant from the Swiss National Foundation (No_163946) to P.E. Jullien. We thank Stefan Grob and Arturo Mari-Ordonez for critical reading of the manuscript.

Received: February 28, 2019
