sRNAs of Plants

Small, Non-coding RNAs of Plants
• Regulatory RNAs that act through gene silencing
• Two classes of small RNAs (sRNAs)
  o microRNA (miRNAs)
    ▪ Encoded by miRNA genes in the genome
  o small interfering RNA (siRNAs)
    ▪ Derived from double stranded RNA produced by multiple genome sources
      • MITES: tomato, rice

• Function
  o Transcriptional silencing
  o Posttranscriptional gene silencing

• Common functional biochemical steps
  1. Double-stranded RNA (dsRNA) molecules are induced
  2. dsRNAs are processed into 18-25 nt sRNA
  3. sRNA is 3’-O methylated
  4. sRNA is incorporated into functional effector complexes

Plant miRNAs
• Encoded by ~100s of MIR genes
  o Low abundance relative to siRNAs
• 20-22 nt in size
  o Contain
    ▪ 2-nt overhangs
    ▪ 5’phosphate and 3’hydroxyl groups
      • On BOTH ends
• Act as regulatory RNAs at the
  o Post transcriptional level
  o Transcriptional level
Synthesis of miRNAs
- Transcribed from nuclear $MIR$ genes by RNA polymerase II
  - Primary transcript forms imperfect fold back structure
    - Called *pri-miRNAs*
      - pri-mRNAs are capped and methylated
- pri-miRNAs converted to
  - *pre-miRNA*
    - folds into stem-loop structure
    - then
      - *miRNA/miRNA*\(^{\ast}\) dsRNA molecule
- Both steps performed in D-bodies
  - D-bodies
    - Nuclear processing center
      - DCL1/HYL1/SE complex
        - DCL1 = dicer
          - RNA folding protein
            - Null mutants are lethal
        - HYL1 = HYPONASTIC LEAVES 1
          - RNA binding function
        - SE = SERRATE
          - C2H2 zinc finger protein
            - Null mutant is lethal
      - pre-miRNA formation aided by
        - Cap-binding proteins
          - CAP80
          - CAP20
- *miRNA/miRNA*\(^{\ast}\) complex methylated at each end
  - Methylation by HEN1
    - Methyltransferase protein
- *miRNA/miRNA*\(^{\ast}\) complex moves to cytosol
  - Movement through the HST1 protein
    - HASTY1 protein
Functional regulation of gene expression

- Process called Post Transcriptional Gene Silencing
  - PTGS
- Functional unit
  - Called RNA-Induced Silencing Complex
    - RISC
- miRNA incorporated into ARGONAUTE 1 (AGO1) protein
  - Arabidopsis AGO family
    - 10 family members
- miRNA* stand degraded
- AGO1/miRNA action
  - Complex binds to exact or nearly-exact target mRNA
    - Binding by base pair complementarity
- AGO1 functions
  - Slices the mRNA
  - OR
    - Binding to mRNA prevents translation
- AGO1/miRNA functions aided by
  - SQUINT
  - HSP90

![Bar chart](image)

### A  No. of miRNA genes

<table>
<thead>
<tr>
<th>Species</th>
<th>Total miRNA genes</th>
<th>TE-like</th>
<th>Non-TE-like</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arabidopsis</td>
<td>154</td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>Papaya</td>
<td>65</td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>Poplar</td>
<td>130</td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>Medicago</td>
<td>122</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Scybean</td>
<td>194</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>Grape</td>
<td>131</td>
<td></td>
<td>115</td>
</tr>
<tr>
<td>Rice</td>
<td>207</td>
<td></td>
<td>171</td>
</tr>
<tr>
<td>Sorghum</td>
<td>110</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Maize</td>
<td>120</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Moss</td>
<td>196</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Green algae</td>
<td>5</td>
<td></td>
<td>0</td>
</tr>
</tbody>
</table>

### B  No of miRNA gene families

<table>
<thead>
<tr>
<th>Species</th>
<th>Total miRNA gene families</th>
<th>TE-like</th>
<th>Non-TE-like</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arabidopsis</td>
<td>77</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>Papaya</td>
<td>20</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>Poplar</td>
<td>21</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Medicago</td>
<td>28</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Scybean</td>
<td>22</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Grape</td>
<td>38</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Rice</td>
<td>56</td>
<td></td>
<td>12</td>
</tr>
<tr>
<td>Sorghum</td>
<td>24</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Maize</td>
<td>21</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Moss</td>
<td>79</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Green algae</td>
<td>5</td>
<td></td>
<td>0</td>
</tr>
</tbody>
</table>
Fig. 1. Biogenesis and function of miRNAs and siRNAs. (A) MIR genes are initially transcribed by PolII into a single-stranded RNA that folds back to form a hairpin structure (also called pri-miRNA) thought to be stabilized by the RNA-binding protein DAWDLE (DDL) [138]. Splicing and further processing in nuclear dicing bodies involve the interactive functions of HYL1 and SE and of the cap-binding proteins (CBP) CBP20 and CBP80 [139] and [140]. Pri-miRNAs and pre-miRNAs are generally thought to be processed from the free-end opposite to the loop by DCL1 to yield one or several phased miRNA/miRNA* duplexes. These are then methylated by HEN1 and transported to the cytoplasm by HST1. The miRNA guide strand is selected, incorporated, and stabilized in dedicated AGO1 protein. miRNA-guided AGO1-containing RISC directs mRNA cleavage or translation inhibition of the target transcript. miRNA-guided AGO1 functions are promoted by SQUINT (SQN) and HSP90 [22] and [141]. (B) Various sources of dsRNA, its processing into siRNAs by one of four DCLs proteins assisted by dsRNA-binding proteins, HEN1-mediated siRNA stabilization, and selected strands of siRNA duplexes guide AGO-containing RISC to target RNAs for endonucleolytic cleavage and for translation repression, or these siRNAs can then guide AGO4 or AGO6 to function in RNA-directed DNA methylation (RdDM) pathway involving PolIV and PolV. Cytosine methylation at these specific sites involves different effectors like the de novo methyltransferase DOMAINS REARRANGED METHYLTRANSFERASE 2 (DRM2); DEFECTIVE IN RNADIRECTED DNA METHYLATION 1 (DRD1), a member of the SWI2-SNF2 chromatin remodeling protein family; and DEFECTIVE IN MERISTEM SILENCING 3 (DMS3), a structural maintenance of chromosomes (SMC) protein.
PTGS via phasiRNAs

- Plant phasiRNAs another plant miRNA type
- Phased small interfering RNAs in plants
- Control expression via PTGS
- tasiRNAs
  - 21-nt *trans*-acting siRNAs
    - One type of plant phasiRNA
      - Occur at 21-26 nt intervals in coding loci
    - phasiRNA loci
    - TAS loci in Arabidopsis
      - Coding loci are source of the phasiRNA
        - NB-LRR, Pentatricopeptide repeat, MYB transcription factor genes
    - PHAS loci in Rice
      - Non-coding are loci source of the phasiRNA
        - Non-coding RNAs transcribed from the locus
    - tsaiRNA generated by cleavage of RNA by AGO1/7 miRNA complex
      - Multiple tsaiRNAs produced from the original transcript
      - Processed 21-nt tsaiRNA then cleaves target mRNAs
    - Gene families regulated by tsaiRNA PTGS
      - Auxin response factor family
      - Pentatricopeptide repeat family
      - MYB transcription factors
  - One function
    - Regulate transition of leaves from juvenile to adult vegetative stage
Fig. 1 A model for phasi/tasiRNA biogenesis in gymnosperms, dicots, and monocots. In plant phasi/tasiRNA biogenesis, (1) phasiRNA precursor transcripts are cleaved by 22-nt miRNA, (2) (RDR)-dependent dsRNAs are synthesized, (3) DCL-dependent processing results in phasiRNAs and (4) their loading onto AGO proteins. The phasiRNA pathway is conserved in gymnosperms, dicots, and monocots. In gymnosperms, there are miR482/miR2118-triggered phasiRNAs derived from both coding-RNAs (NB-LRRs, etc.) and non-coding RNAs (reproductive lincRNAs). The phasiRNA pathway, involving miR482/miR2118-triggered coding-RNAs (NB-LRRs), is also conserved in dicots. In contrast, miR482/miR2118-triggered phasiRNAs generated from reproductive lincRNAs are conserved in monocots.
Plant siRNAs
- Most abundant sRNAs
- 21-26 nt long nucleotides
  - Derived from double stranded RNA (dsRNA) molecules
- Regulate gene expression by
  - Post-transcriptional gene silencing (PTGS)
    - Slicing mRNAs
    - Similar to miRNAs
  - Transcriptional gene silencing (TGS)
    - Methylation of mRNAs

How are siRNAs generated?
- Step 1: dsRNAs molecules production
  - Transcription of inverted repeat loci
  - Replication of virus particles
  - Transposons
  - RNA-dependent RNA polymerase (RDR2) activity
    - Uses single stranded RNA molecules as template
  - Anti-sense RNAs
    - Introduced as a transgene
- Step 2: production of siRNA duplexes
  - Produced by RNA polymerase IV
    - Uses dsRNA as template
  - Complexed with
    - CLASSY1 (CSLY)
      - SNF2-like chromatin remodeling factor
- Step 3: cleavage of dsRNAs into siRNAs
  - Complex of proteins
    - DCL1-4
      - Dicer protein that cleaves the dsRNA
    - DRB4
      - Double stranded binding protein
    - HEN1
      - Methylates siRNA
Post Transcription Gene Slicing by siRNA
- Similar function as miRNA
  - Form RISC complex with AGO1
    - mRNA is sliced
    - Translation is inhibited by binding of AGO1/siRNA complex

Transcriptional Gene Silencing
- RNA directed DNA methylation
  - RdDM
- siRNA complexes with AGO4 or AGO6
  - siRNA/AGO complex interacts with
    - RNA dependent RNA polymerase V (PolV)
    - DDR complex consisting of
      - DRD1
        - SNF2-like chromatin remodeling factor
      - DMS3
        - Maintains chromosome structure
      - RDM1
        - Methylase
      - Novel protein
        - Unknown protein
    - Produces RNAs that interact with siRNAs
      - siRNA/AGO and PolV+DDR complex
    - siRNA/AGO + PolV+DDR + siRNA target RNA
      - Recruits methylation and histone modifying factors to target locus
        - Target locus is modified
Transposons, repeats, etc. → Pol IV + CLSY1 → ss RNA Pol IV transcript → RDR2 → dsRNA → DCL3 → 24 nt siRNA → AGO4-bound siRNA complex + Pol V + DDR Complex → Target Locus → Nascent RNA or Recruitment of complexes for de novo DNA methylation (DRM2) and histone modifications

Spreading the primary methylation signal
  • Reinforces existing heterochromatic regions


**Figure 2.** RNA polymerase IV-dependent production of small RNAs (1° siRNAs) reinforces existing heterochromatic regions by primary RNA-directed DNA methylation (1° RdDM). Primary RdDM can lead to the production of 2° siRNAs which trigger the spreading of methylation into adjacent regions, resulting in 2° RdDM. Secondary RdDM results from a Pol IV-derived aberrant RNA transcribed from methylated target templates. Some of these sRNAs may act in *trans* to direct 2° RdDM in a Pol V-dependent manner.
sRNA Nomenclature

- Consistent nomenclature used across species
    - Based on sequence of hairpin precursor
  - Revision to miRNA annotation
      - Correct errors that called siRNAs as miRNAs
        - Calls for data based on whole genome sequencing of the small RNA fraction in replicate
- [http://www.mirbase.org/](http://www.mirbase.org/)
- Release 22: March 2018
  - Major update
    - Release 21: June 2014
- 38,589 hairpin precursors
  - 10,031 new entries since
- 48,885 mature miRNAs
  - 13,149 new entries since
- 271 species
  - 43 new species since
- Example
  - miR166
    - 51 species
    - 299 hairpin sequences
    - 369 mature miRNAs
sRNA Regulation

- sRNA targets
  - Transcription factors major target of sRNAs
  - Biochemical pathway genes are also targets
- All biological types of pathways affected by siRNAs
  - sRNAs can be involved in multiple pathways
- Response to biological process
  - sRNA induction
    - Reduces expression of target mRNA
  - sRNA repression
    - Increases expression of target mRNA
Fig. 2. Summary of stress-regulated small RNAs and their target families. Small RNAs are categorized based on the stress that they respond to. Grey box: abiotic stress, dark grey: biotic stress, green boxes: upregulated small RNA, light green boxes: slightly upregulated small RNA, and red boxes: downregulated small RNA. At (Arabidopsis thaliana), Bd (Brachypodium distachyon), Bn (Brassica napus), Br (Brassica rapa), Gm (Glycine max), Mt (Medicago truncatula), Nt (Nicotiana tabacum), Os (Oryza sativa), Pt (Populus trichocarpa), Pta (Pinus taeda), Ptr (Populus tremula), Pp (Physcomitrella patens), Pv (Phaseolus vulgaris), Ta (Triticum aestivum), Zm (Zea mays). See the text for references.
Fig. 3. Regulatory network of stress-responsive miRNAs in Arabidopsis. A network is proposed that describes the molecular mechanisms underlying the response of Arabidopsis plants to different biotic and abiotic stresses. The network is based on the changes in expression profiles of miRNA and subsequent target transcripts in plants under stress. **Green boxes:** upregulated RNAs; **red boxes:** downregulated RNAs.
Examples

Disease Resistance

General miRNA Model of Disease Resistance


- **Normal, non-infection condition**
  - miRNAs cleave NLR transcripts
  - Resistance gene not active
  - Secondary phasiRNA cleavage products regulate downstream genes activated upon infection

- **Disease challenge condition**
  - miRNA production down regulated
  - Disease resistance system turned on
Recent In-depth Analysis


Nucleotide Leucine-Rich Genes (NLRs)

- Mediate many disease resistance responses
- Two structural classes
  - Toll-interleukin receptor – Nucleotide Binding Site – Leucine-rich repeats
    - TLR abbreviation
  - Coiled-coiled domain – Nucleotide Binding Site – Leucine-rich repeats
    - CLR abbreviation
- Tobacco NLR Genes
  - 624 NLR genes
  - 25% NLR gene targeted by miRNAs
  - 98% NLR genes generated phasiRNAs
  - 210 sRNA silencer (miRNAs + phasiRNAs)
- Tomato NLR Genes
  - 177 NLR genes
  - 50% of NLR genes targeted by miRNAs
  - 96% NLR genes generated phasiRNAs
  - 747 sRNA silencer (miRNAs + phasiRNAs)
- Majority of NLR genes regulated by siRNAs during plant growth
  - sRNAs specific to TLR and CLR classes
  - phasiRNAs have a greater impact on NLR regulation
- Pattern during growth from 1 week to 3 and 6 weeks
  - sRNA levels decreased from 1 week to 3 and 6 weeks
  - NLR transcript levels increased