


# Plant Non-coding RNAs

- **General background and many details of the topic from:**
  - Yu et al. 2019. Plant Noncoding RNAs: Hidden Players in Development and Stress Responses. Annual Review of Cell and Developmental Biology 39:407
  - Rai et al. 2019. Classification and experimental identification of plant long non-codingRNAs. Genomics 111:997.

## The Plant Transcriptome Landscape Puzzle

- Plant genomes very greatly in size
  -  **0.1 Gb - 25.0 Gb**
- Most species rich in TE, repeats, and satellite DNA
- Early on this DNA was considered Junk DNA
  - **Junk DNA** = non-functional (protein coding) DNA
- But 90% of the entire genome is transcribed
  - But only 2% of RNA used to encode protein
- **What are the other RNAs?**
- **What is the function/relevance of the other RNA products?**

Major questions that are addressed by studying non-coding RNA.

## Transcription Products Known for A Long Time

- **mRNA for Protein Coding Genes**
  - Functions are to encode
  - Enzymatic, structural, regulatory functions
  - Responsible for nearly all phenotypes
- **Traditional non-coding RNA**
  - tRNAs = Transfer RNAs
  - rRNAs = Ribosomal RNAs
  - snRNAs = Small nuclear RNA
    - Processing of hnRNA
  - snoRNAs = Small nucleolar
    - Involved in chemical modification of other RNAs
      - rRNAs, tRNAs, snRNAs
        - Methylation
        - Pseudouridylation

What was known before genomic tools were developed.

## Non-coding RNAs: Genomic Era Discovery

### Major features of plant ncRNAs

- Provide no (or very little) protein coding function
- Act as regulatory elements
- Primarily located in the intergenic (between genes) regions of the genome

## General Classes of Plant Non-coding (nc) RNAs

- **Small RNAs**
  - 18-30 nucleotides (nt) in length
- **Medium-sized ncRNAs**
  - 31-200 nt in length
- **Long ncRNA (lncRNA)**
  - >200 nt in length

# Small, Non-coding RNAs of Plants

- **Regulatory RNAs that act through gene silencing**

- **Function via**

- Transcriptional silencing
- Posttranscriptional gene silencing

How sRNAs function.

- Three classes of small RNAs (sRNAs)

- **microRNA (miRNA)**

- Encoded by miRNA genes in the genome
- Typically, 21nt long
- Regulate expression by
  - Cutting mRNA transcript or repressing translation

- **small interfering RNA (siRNA)**

- Derived from double stranded RNA produced by transposable elements and repeat elements
- Typically, 24nt long
- Regulate expression by
  - DNA methylation OR histone methylation

- **Phased small interfering RNAs (phasiRNA)**

- Derived from
  - Cleavage products of mRNA
    - Dicots
      - Nucleotide-binding leucine-rice repeat (NLR) encoding genes
      - Pentatricopeptide repeat (PRR) encoding genes
    - Monocots
      - Long-non coding RNAs (lncRNA)
  - Non-coding trans-acting siRNAs TAS transcripts (tasiRNA)
- Typically 21-24nt long

- **Common miRNA, siRNA, and phasiRNA functional biochemical steps**

1. Double-stranded RNA (dsRNA) molecules are induced
2. dsRNAs are processed into 18-25 nt sRNA
3. sRNA is incorporated into functional effector **AGO (Argonaute)** complexes that affect gene expression

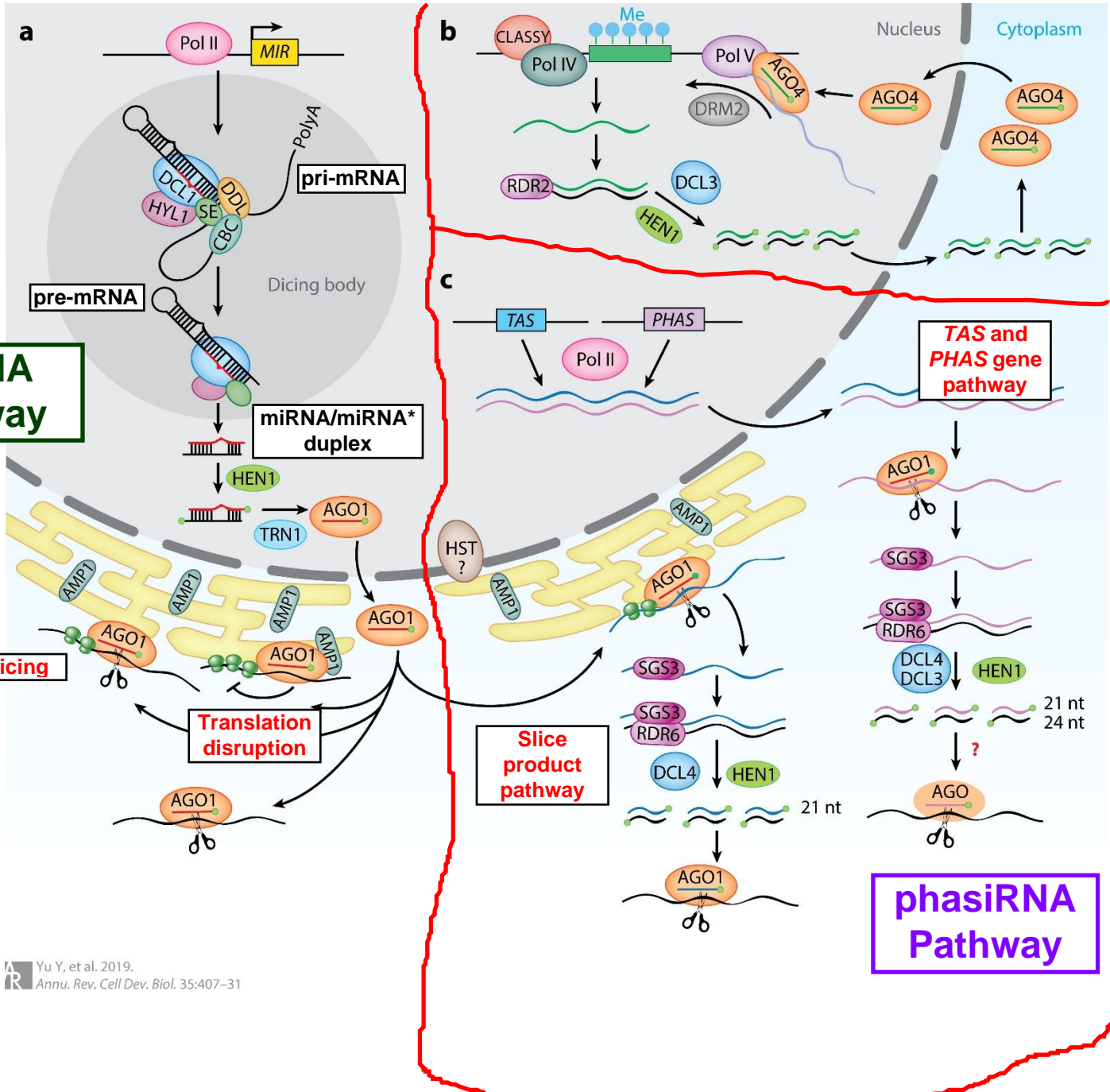
Classes (and general features) of sRNAs!!!

Common sRNAs features.

# Synthesis and Evolution of Plant Small RNAs

- Yu et al. 2019. Ann Rev Cell Develop Biol 35:407

**siRNA Pathway**



**Figure 1** Biogenesis and modes of action of plant small RNAs. **(a)** A **MIR gene is transcribed** into a pri-miRNA, which is sequentially **processed** first **into** a pre-miRNA and then into a **miRNA/miRNA\* duplex**. The **duplex** is **methylated** by HEN1, and the miRNA strand is **loaded into AGO1** in the nucleus. The **miRNA-AGO1** complex is **transported** to the **cytoplasm** and **regulates target gene** expression through **transcript cleavage and/or translation repression**. **(b)** Pol IV generates **single-stranded siRNA precursors**, which are **converted into dsRNAs** and **processed** into **24-nt siRNA duplexes**. **Methylated** siRNAs are **loaded into AGO4** in the **cytoplasm** and are **transported** to the nucleus, followed by the recruitment of these **siRNA-AGO4 complexes to Pol V transcripts**. The subsequent **recruitment of DRM2 catalyzes DNA methylation at RdDM target loci**. **(c)** **TAS or PHAS loci** are **transcribed** into **single-stranded RNAs** that are **targeted** by an **miRNA-AGO1/7 complex**. The 5' or 3' **cleavage fragment** is **protected** by SGS3 and **converted** into **dsRNA** by RDR6. DCL proteins **process** these dsRNAs into **21- or 24-nt phasiRNAs**. The 21-nt tasiRNAs, which are phasiRNAs from TAS loci, are primarily **loaded** into **AGO1** and guide **transcript cleavage** of their targets.

**Abbreviations:** **AGO**, ARGONAUTE; **AMP1**, ALTERED MERISTEM PROGRAM 1; **CBC**, CAP-BINDING COMPLEX; **DCL**, DICER-LIKE; **DDL**, DAWDLE; **DRM2**, DOMAINS REARRANGED METHYLASE 2; **dsRNA**, double-stranded RNA; **HEN1**, HUA ENHANCER 1; **HST**, HASTY; **HYL1**, HYPONASTIC LEAVES 1; **Me**, methylated; **phasiRNA**, phased siRNA; **Pol**, RNA polymerase; **pre-miRNA**, precursor miRNA; **pri-miRNA**, primary miRNA; **RdDM**, RNA-directed DNA methylation; **RDR2/6**, RNA-DEPENDENT RNA POLYMERASE 2/6; **SE**, SERRATE; **SGS3**, SUPPRESSOR OF GENE SILENCING 3; **siRNA**, small interfering RNA; **tasiRNA**, trans-acting siRNA; **TRN1**, TRANSPORTIN 1.

## Plant miRNAs

- Encoded by ~100s of **MIR** genes
  - Low abundance relative to siRNAs
- 20-22 nt in size
  - Contain
    - 2-nt overhangs
    - 5'phosphate and 3'hydroxyl groups
      - On BOTH ends
- Act as regulatory RNAs

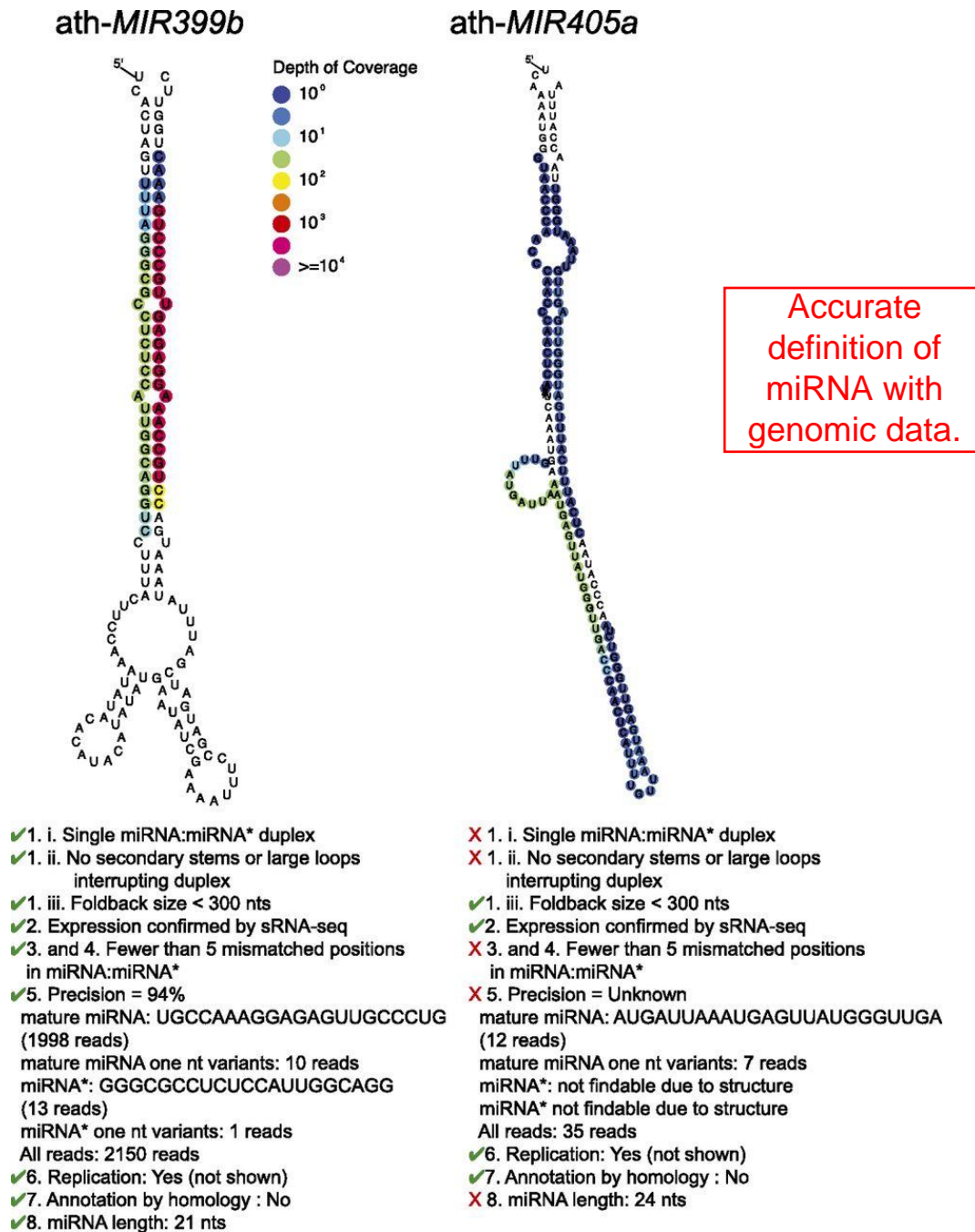
## Defining Plant MicroRNAs

- **Axtell and Meyers 2018. Plant Cell 30:272.**
- **Requirements for acceptance by the Plant MicroRNA Database (miRbase)**
  - <http://www.mirbase.org/>

**Table 1. Updated Criteria for Plant miRNA Annotations**

	2008 Criteria	2018 Criteria
1	One or more miRNA/miRNA* duplexes with two-nucleotide 3' overhangs	Add requirements that exclude secondary stems or large loops in the miRNA/miRNA* duplex and limit precursor length to 300 nucleotides
2	Confirmation of both the mature miRNA and its miRNA*	Disallow confirmation by blot; sRNA-seq only
3	miRNA/miRNA* duplex contains ≤4 mismatched bases	Up to five mismatched positions, only three of which are nucleotides in asymmetric bulges
4	The duplex has at most one asymmetric bulge containing at most two bulged nucleotides	Up to five mismatched positions, only three of which are nucleotides in asymmetric bulges
5	≥75% of reads from exact miRNA or miRNA*	Include one-nucleotide positional variants of miRNA and miRNA* when calculating precision
6	Replication suggested but not required	Required; novel annotations should meet all criteria in at least two sRNA-seq libraries (biological replicates)
7	Homologs, orthologs, and paralogs can be annotated without expression data, provided all criteria met for at least one locus in at least one species	Homology-based annotations should be noted as provisional, pending actual fulfillment of all criteria by sRNA-seq
8	miRNA length not an explicit consideration	<b>No RNAs &lt;20 nucleotide or &gt;24 nucleotides should be annotated as miRNAs. Annotations of 23- or 24-nucleotide miRNAs require extremely strong evidence.</b>

# Examples of 2018 Valid/Invalid Plant miRNAs



**Figure 1.**

Examples of Valid and Invalid miRNA Loci. Small RNA-seq data were from GSE105262 (Polydore and Axtell, 2017). Left: Arabidopsis (ath) MIR399b, a locus judged valid under the new criteria. Right: ath-MIR405a, a locus judged invalid under the new criteria



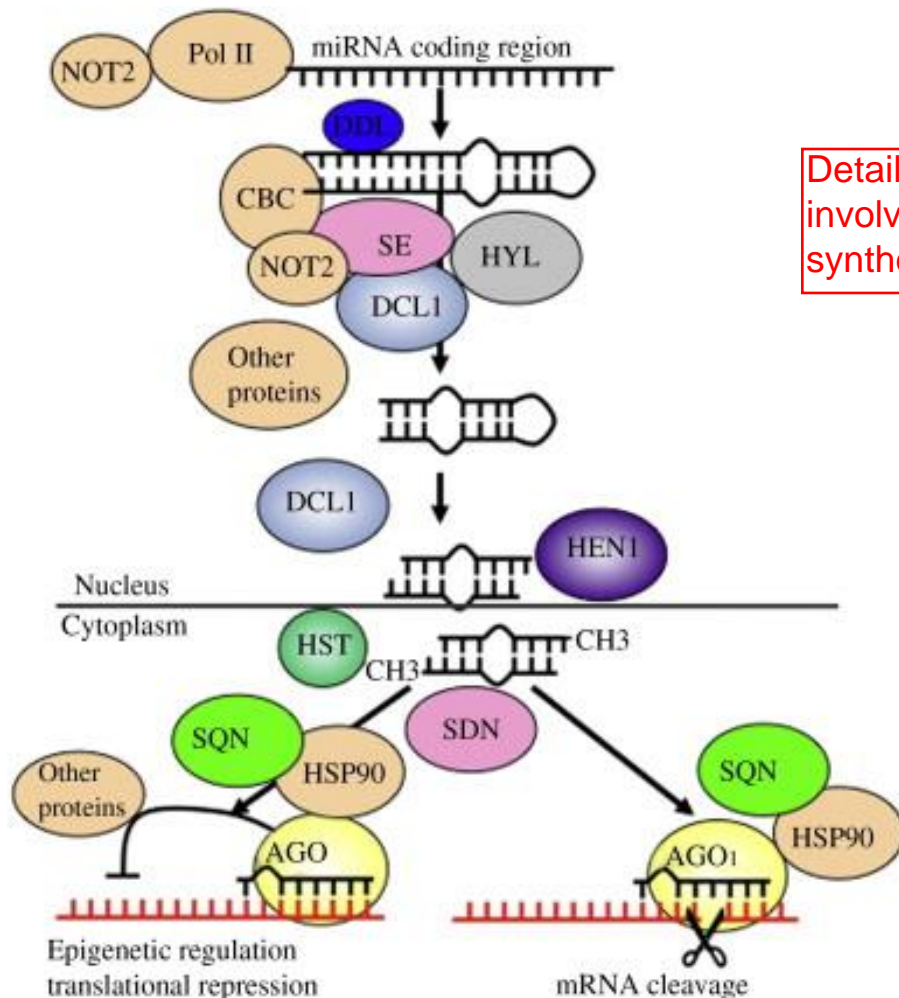
# Number of miRNAs in Plants and Animals

- Data from miRbase database (<http://www.mirbase.org/>)
- April 20, 2020

Common name	Genus/species	# miRNA primary transcript	# miRNAs
Grape	<i>Vitis vinifera</i>	159	178
Arabidopsis	<i>Arabidopsis thaliana</i>	326	428
Corn	<i>Zea mays</i>	172	325
Rice	<i>Oryza sativa</i>	588	654
Mosquito	<i>Aedes aegypti</i>	155	192
Nematode	<i>Caenorhabditis elegans</i>	253	450
Drosophila	<i>Drosophila melanogaster</i>	267	495
Chimpanzee	<i>Pan troglodytes</i>	697	707
Human	<i>Homo sapiens</i>	918	2883

## Functional Components of Plant miRNAs

- Wu 2013. J of Genetics and Genomics 40:217



Details of complexes involved in miRNA synthesis function.



## Synthesis of miRNAs

### 1. 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

### 2. *pri-miRNAs* converted to

- *pre-miRNA*
  - folds into stem-loop structure
    - then
  - *miRNA/miRNA\** dsRNA molecule formed
- 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

### 3. *miRNA/miRNA\** complex methylated at each end

- **Methylation by HEN1**
  - Methyltransferase protein

### 4. *miRNA/miRNA\** complex moves to cytosol

- **Movement through the HST1 protein**
  - HASTY1 protein

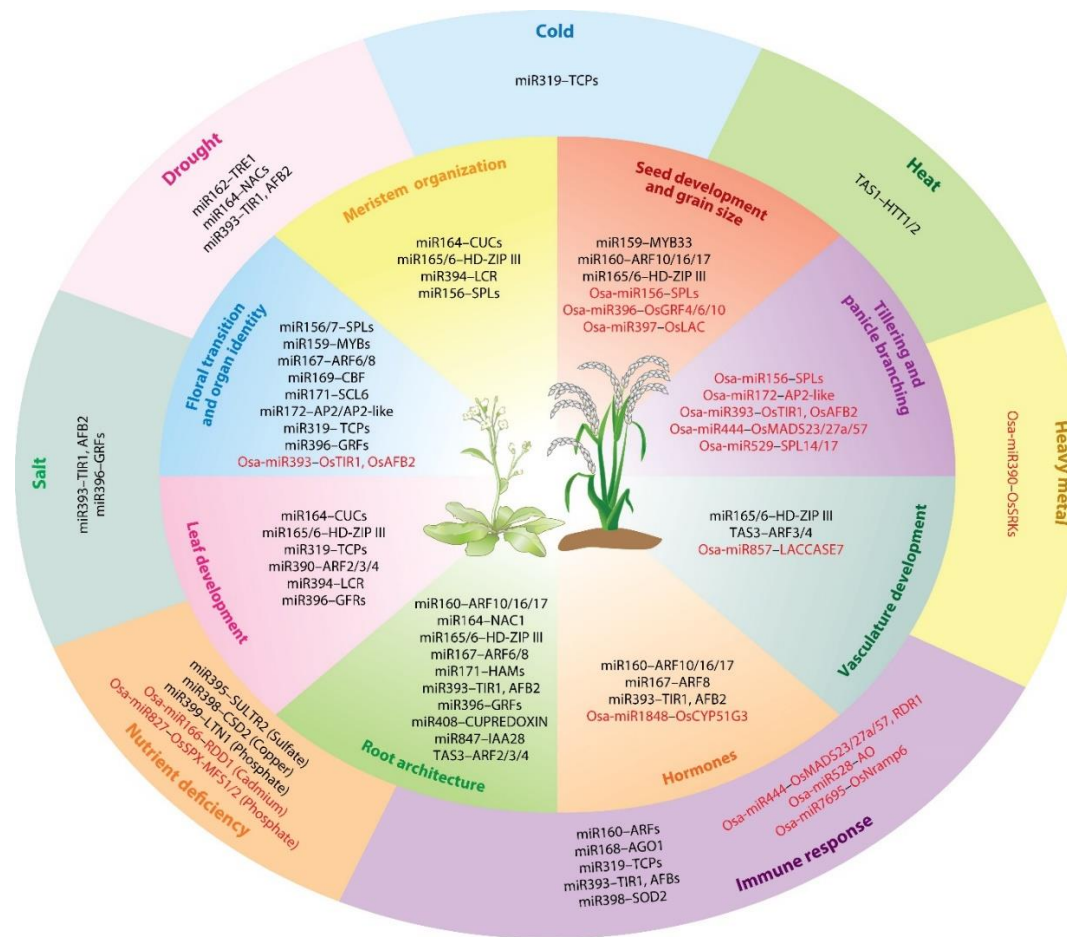
## Functional regulation of gene expression by miRNAs

- 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\* strand degraded
- **AGO1/miRNA complex action**
  - **Complex binds to exact or nearly-exact target mRNA**
    - Binding by base pair complementarity
  - **AGO1 functions by**
    - **Slice at the phosphodiester bond between nucleotides 10 and 11 of the miRNA**
      - OR
    - **Binding to mRNA prevents translation**
- AGO1/miRNA functions aided by
  - SQUINT
  - HSP90

**RISC  
Complex  
Performs  
slicing  
function**

# Plant Biological Processes Controlled by miRNAs

- Yu et al. 2019. Plant Noncoding RNAs: Hidden Players in Development and Stress Responses. *Ann Rev Cell Develop Biol* 35:407



**miRNAs involved in ALL stages of plant growth, development, and abiotic/biotic stress response!!!**

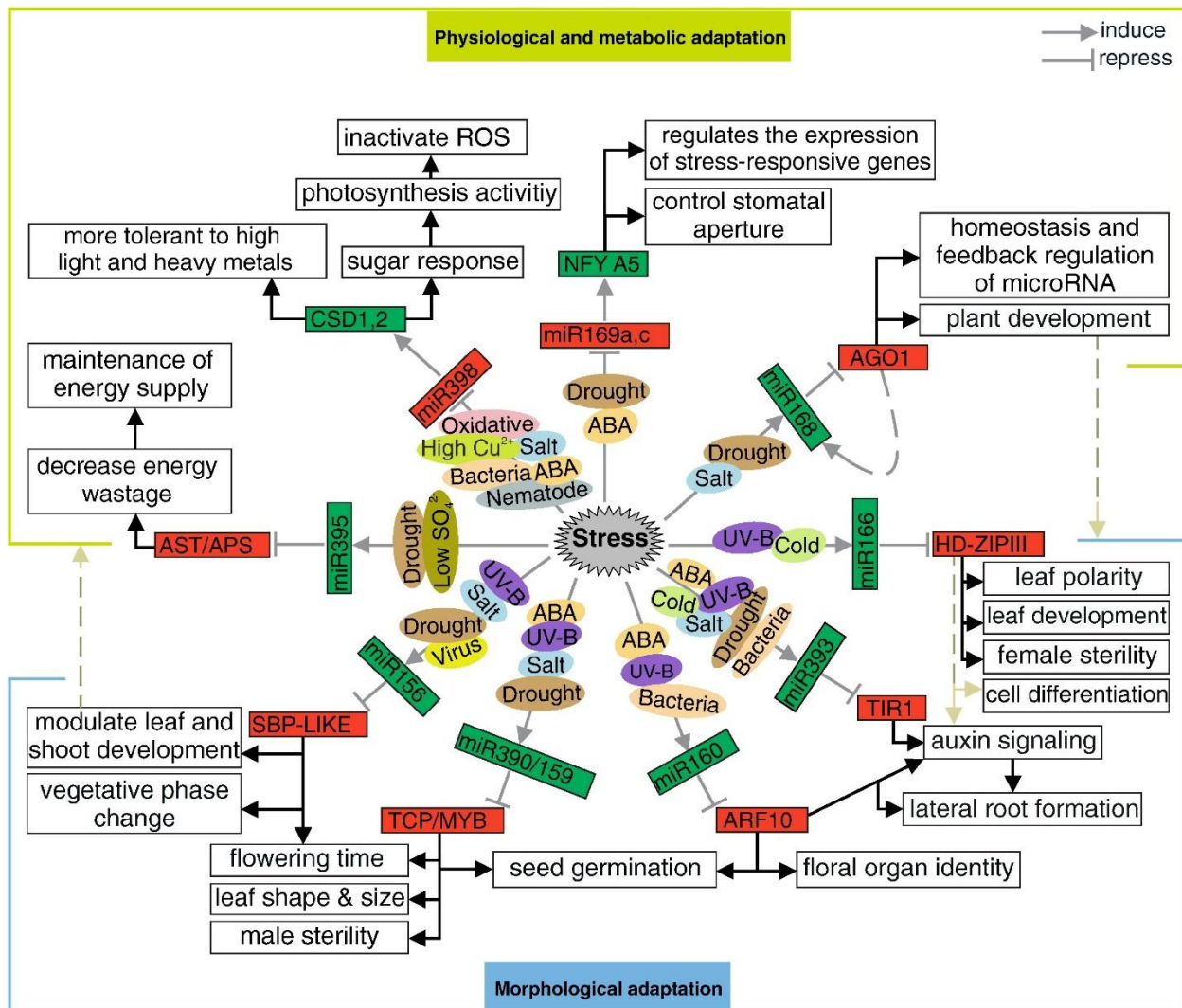
Yu Y, et al. 2019. *Annu. Rev. Cell Dev. Biol.* 35:407-31

**Figure 2** Functions of miRNAs in plant development and stress responses and an overview of the current understanding of miRNA-mediated regulation during development (inner circle) and responses to biotic and abiotic stresses (outer circle) in Arabidopsis and rice. Red font indicates miRNA-target modules that act in rice, but not in Arabidopsis.



# Plant miRNA Stress Regulatory Pathway

- Khraiwesh et al. 2012 Biochimica et Biophysica Acta 1819:137



**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.

**Many signals  
activate miRNAs  
to regulate plant  
growth and  
development!!!**



## Plant siRNAs Background

- Derived (mostly) from
  - Heterochromatin repeats
  - Transposable elements
- Modify gene expression by
  - DNA methylation
  - Histone methylation
- Process called
  - RNA-directed DNA methylation (RdDM)



# Maintenance of Heterochromatin Methylation by siRNA

- Source sequences
  - TEs and repeats
    - **Goat: Add methyl groups to source loci**
- Steps of siRNA production
  - 1. Creation of siRNA precursors**
    - **Pol IV** recruited by CLASSY remodeling factors
    - Move to RdDM target loci
    - Create single-stranded siRNA precursor molecules
  - 2. Synthesis of functionally mature 24-nt siRNAs complex**
    - Double stranded RNA (dsRNA) produced by RNA-DEPENDENT RNA POLYMERASE 2 (**RDR2**)
    - dsRNA processed by DCL3 into 24-nt siRNA
    - 24-nt siRNA loaded into AGO4 (Argonaut 4)
  - 3. Meanwhile,**
    - Noncoding transcripts at the same RdDM loci produced
      - Product of RNA polymerase V
        - A plant-specific RNA polymerase
  - 4. Formation of the methylation complex**
    - Non-coding transcript complexes with 24nt siRNA/AGO4 complex
    - DRM2 (DNA REARRANGED METHYLASE 2) recruited to the complex
    - Methylation state established at RdDM loci by action of DRM2
- **Additional Features**
  - A. Histone H3 lysine 9 methylation and DNA methylation
    - A repressive state
      - Assists in recruiting Poly IV (for siRNA production)
  - B. Maintains methylation at CG, CHG, and CHH sequences in
    - **Heterochromatin**
    - **TEs**

Maintaining heterochromatin methylation important for chromosomal integrity from generation to generation!!!

## **Phasi-RNAs**

- Unique feature of Phasi-RNA
  - Source nucleic acid
    - mRNA fragments from miRNA slicing
      - Cleavage products of miRNA RISC complex action
    - TAS (Trans-acting siRNA) loci

### **Steps of Phasi-RNA production**

#### **1. PHAS or TAS loci transcribed**

- **Loci cleaved by AGO1/7 RISC complex**
- Product bound on 5' end by SGS3 protein
- SGS3-bound nucleotide converted into double-stranded RNA
  - Conversion to dsRNA mediated by RDR6

#### **2. dsRNAs cleaved**

- By DCL proteins

#### **3. Cleavage product is methylated**

- By HEN1
- Methylated product loaded into AGO1 complex

#### **4. AGO1 loaded complex cleaves target loci**

#### **Phasi-RNA:**

**A method to control a genes expression by using slicing products of its mRNA pool to slice other mRNA in its pool.**

# Unique Functions of Phasi-RNAs in Regulating Disease Defense in Dicots

- **Plant Resistance Associated Genes**
  - **Nucleotide-binding Site, Leucine Rice Repeat Sequences**
    - NB-LRR genes
      - Exists as gene families
        - NB-LRR largest class of genes that PhasiRNAs
      - NB-LRR transcripts
        - Cleaved by 22-nt miRNAs
          - Generate
            - 21-nt DCL4-dependent phasiRNAs
      - NB-LRR derived phasiRNAs
        - Post-transcriptionally regulate NB-LRR expression
  - **Pentatricopeptide repeat (PPR) gene family**
    - Some members regulated by tasiRNAs
      - Transcripts cleaved by miRNAs
    - Derived phasiRNAs of PPR
      - Silence pathogen transcripts

# Plant Long Non-coding RNA (lncRNA)

## Types of Plant lncRNAs

- **Linear**

- Derived from:

- Intergenic regions

- Long intergenic ncRNAs = **lincRNAs**

- Introns

- Intronic ncRNAs = **incRNAs**

- Coding regions

- Natural antisense transcripts = **lncNATs**

- **Circular**

- Derived from:

- Coding or intronic regions

- Circular RNAs = **circRNAs**

## Linear lncRNAs

- Synthesized by

- **RNA Polymerase II**

- **Show typical mRNA features**

- 5' methylated CAP
- PolyA+ tail

- **Biochemical roles**

- As RNA molecule with other RNAs or proteins
- Not Translated

## lncNATs

- Expression CORRELATED with level of sense strand mRNA expression

- 70% of Arabidopsis mRNAs have associated lncNATS

- **lncNATS associated with histone methylation**

- Components of scaffolds that recruit
  - DNA-modifying and histone-modifying enzymes

- **Induce epigenetic changes that**

- **Regulate (up or down) gene expression**

**Anti-sense  
self-control of  
a gene**

**Multiple  
sources of  
lncRNAs in  
eukaryotic  
organisms!!!**

## lincRNAs

- Expression NOT CORRELATED with neighboring gene expression
- Function in trans
  - Target gene elsewhere in the genome
- Example:
  - lincRNA ELENA1
    - Interacts with transcription module
      - Activates pathogen resistance protein PR1

## incRNA Example

- COLDAIR incRNA
  - Transcribed from first intron of flowering gene *FLC*
  - Interacts with lincRNA COLDWARP a *FLC* promoter lincRNA
    - *FLC* gene expression is silenced

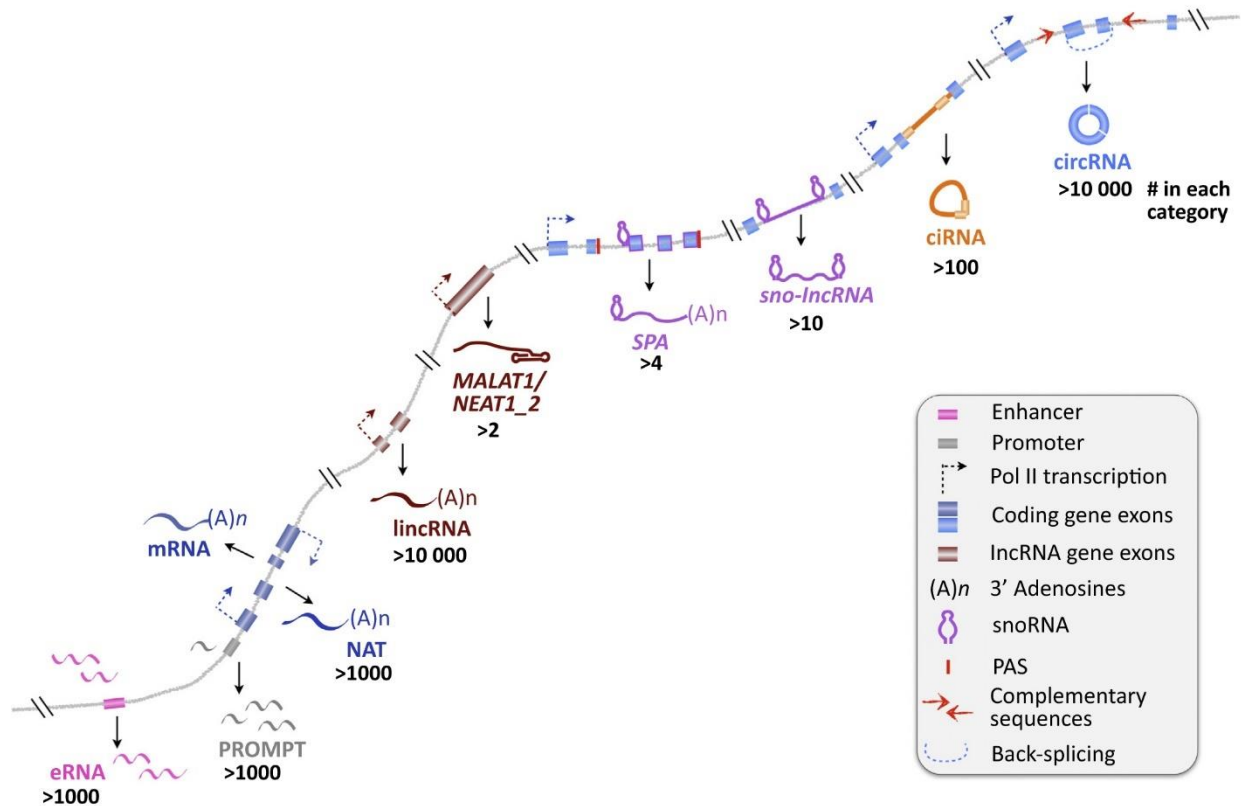
Intron sequeces  
controls its own  
gene's expression!!!

## Circular lincRNAs

- Non-polyadenylated
- Synthesized by
  - Back-splicing of internal exons
    - Exported to cytoplasm
  - Intron lariats
    - Nuclear localized
- Few functions discovered

# Cartoon of All Human lncRNAs

- From: Yang and Chen. 2017. Trends in Genetics 33:540



**Figure 1.** A schematic drawing to illustrate the diversity of lncRNAs in mammalian genomes. **ciRNA**, circular intronic RNA derived from intron lariats; **circRNA**, circular RNA produced from back-splicing of exons; **eRNA**, enhancer RNA; **lincRNA**, large intervening/intergenic noncoding RNA; **NAT**, natural antisense transcript; **PROMPT**, promoter upstream transcript; **sno-lncRNA**, small nucleolar RNA (snoRNA)-ended lncRNA; **SPA**, 50 snoRNA-ended and 30-polyadenylated lncRNA; metastasis-associated lung adenocarcinoma transcript 1 (**MALAT1**)/Nuclear enriched abundant transcript 1, the long isoform (**NEAT1\_2**), lncRNA with the 30 end alternatively processed by ribonuclease P (RNase P). The number of lncRNAs identified in each category in mammals is shown beneath



# Functions of Plant lncRNAs

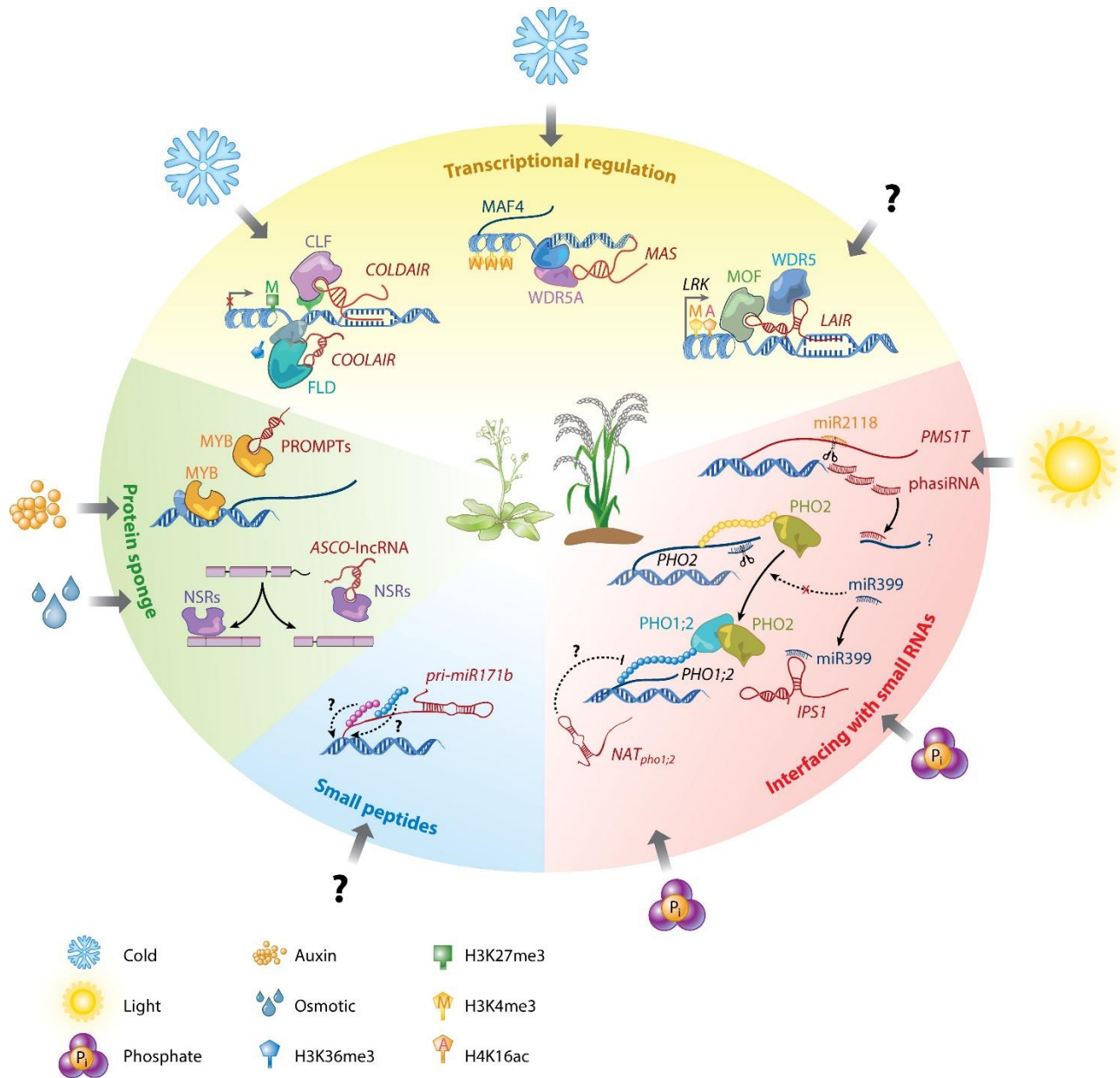
- Table from: Rai et al 2019. *Genomics* 115:997.

**Table 1.** List of plant long noncoding RNAs identified on the basis of their biological function and mode of action.

SPECIES	Name	Mechanism of action	Biological function	References
Arabidopsis thaliana	IPS1	Target mimicry: interacts with an miRNA, ath-mir399	Regulates <b>phosphate homeostasis</b>	[22]
	COLDAIR	<b>Epigenetic silencing</b> : modification of histones	Regulates <b>flowering</b>	[27]
	COOLAIR	<b>Epigenetic silencing</b> : Promoter interference	Regulates <b>flowering</b>	[88]
	HID1	Stage specific: Associates with chromatin and <b>represses transcription</b> of phytohormone interaction factor 3 (PIF3)	Regulates <b>photomorphogenesis</b>	[103]
	asHSFB2a	<b>Hijacks</b> nuclear Antisense transcription regulators	Regulates the <b>gametophytic and vegetative development</b>	[111]
	ASCO-lncRNA	Regulators of <b>alternate splicing</b>	Regulates the <b>development of lateral roots</b>	[5]
	APOLO	Causes the <b>formation of Chromatin loop</b> and regulates the expression of its neighboring genes	Regulated <b>plant development</b> through auxin signaling	[2]
	ELENA1	<b>Increases the expression</b> of genes involved in plant immunity, such as PR1 and PR2, B 1,3glucanase and salicylic acid genes	Upregulated in <b>disease</b> caused by Pseudomonas syringae	[61]
	Drought induced RNA (DRIR)	<b>Regulates</b> stress responses such as stomata closure	Upregulated in <b>drought and salt stress</b>	[72]
	CDF5 LONG NONCODING RNA (FLORE)	<b>Natural anti-sense transcript</b> of CDF5, upregulates flowering	<b>Photomorphogenesis</b>	[26]
Glycine max (soybean)	GmENOD40	Involved in the development of nodules	Regulates the <b>re-localization of proteins</b> from nucleus to cytoplasm	[117]
Hordeum vulgare (Barley)	HvCesA6 lnc-NAT	Acts as a <b>precursor for siRNA</b>	Involved in the <b>synthesis of cell wall</b>	[25]
Medicago truncatula	MtENOD40	Involved in the development of nodules	Regulates the re-localization of proteins	[83]
Oryza sativa (Rice)	OsPI1	Unknown	Regulates <b>phosphate homeostasis</b>	[106]
	OsENOD40	Involved in the development of nodules	Unknown	[42]
	LDMAR (P/TMS12-1)	Acts by causing <b>promoter methylation</b>	<b>Male fertility</b> regulation	[17]
	Cis-NATPHO1;2	Acts by <b>enhancing translation</b> of PHOSPHATE1;2	Regulates <b>phosphate homeostasis</b>	[30]
	lncRNAs: 2224 lincRNAs: 1624 lncNATs: 600	Tissue and stage specific: Anthers, pistils, seeds five days after pollination, and shoots 14 days after germination	Regulates <b>sexual reproduction</b>	[128]
Petunia hybrid	SHO lnc-NAT	<b>Degrades dsRNA</b>	Regulates the <b>synthesis of local cytokines</b>	[138]
Populus trichocarpa	lncRNA20 lincRNA2752	<b>target mimic</b> of ptc-miR476 target mimic of ptc-miR169	Upregulated in <b>abiotic stress</b> (drought)	[80]
Solanum lycopersicum	TPS11	Unknown	Regulates <b>phosphate homeostasis</b>	[53]
	Slylnc0195 slylnc1077	<b>target mimic</b> of miR166 target mimic of miR399	Upregulated in infections caused by tomato yellow leaf curl virus	[97,101]

# Functional Roles of lncRNAs

- From: Yu et al. 2019. *Ann Rev Cell Devel Biol* 39:407



Yu Y, et al. 2019. *Annu. Rev. Cell Dev. Biol.* 35:407–31

Figure 3 Representative models for the roles of plant long noncoding RNAs (lncRNAs). **lncRNAs** can **serve as scaffolds, molecular mimics and sponges, and small interfering RNA precursors**. They can also be translated into small peptides to regulate target genes in cis or in trans during plant development. The expression of most reported **plant lncRNAs** is **induced by diverse environmental conditions**. COLDAIR, COOLAIR, MAS, and LAIR are lncRNAs regulating mRNA transcription in cis. PMS1T is a lncRNA acting as a phasiRNA precursor. IPS1 regulates PHO2 by acting as the endogenous target mimic of miR399 and affects phosphate homeostasis together with another lncRNA, NAT<sub>pho1;2</sub>. Pri-miR171b is a peptide-encoding lncRNA. PROMPTS and ASCO-lncRNA are protein-binding lncRNAs that suppress the function of target proteins.