Endogenous Small RNA Clusters in Plants

Yong-Xin Liu 1,2, Meng Wang 1, Xiu-Jie Wang 1,*

1 State Key Laboratory of Plant Genomics, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, China
2 University of Chinese Academy of Sciences, Beijing 100101, China

* Corresponding author.

E-mail address: xjwang@genetics.ac.cn (Wang XJ).

Tel: +86 10 6480 6590

Running title: Liu YX et al / Endogenous Small RNA Clusters in Plants

Abstract

In plants, small RNAs (sRNAs) usually refer to non-coding RNAs (ncRNAs) with lengths of 20~24 nucleotides. sRNAs are involved in the regulation of many essential processes related to plant development and environmental responses. sRNAs in plants are mainly grouped into microRNAs (miRNAs) and small interfering RNAs (siRNAs), and the latter can be further classified into trans-acting siRNAs (ta-siRNAs), repeat-associated siRNAs (ra-siRNAs), natural anti-sense siRNAs (nat-siRNAs), etc. Many sRNAs exhibit clustered distribution pattern in the genome. Here, we summarize the features and functions of cluster-distributed sRNAs, aimed to not only provide a thorough picture of sRNA clusters (SRCs) in plants, but also shed light on the identification of new classes of functional sRNAs.

KEYWORDS: Small RNA cluster; MicroRNA; siRNA; Expression pattern
Introduction

Small RNAs (sRNAs) have been identified as pivotal regulators in plants, participating in the regulation of plant growth and development, adaptation to abiotic stresses and responses to biotic pathogens [1,2]. sRNAs in plants can be mainly classified into microRNAs (miRNAs) and small interfering RNAs (siRNAs) [3]. They are processed from the double-stranded region of single RNA molecules with hairpin-shaped secondary structures or double-stranded RNA (dsRNA) molecules, respectively, by one of the dicer-like family enzymes (DCL) encoded by the plant genome [4].

miRNAs are most extensively studied sRNAs so far, and have been indicated to play essential roles in many aspects of plant development, such as meristem division, organ specification, leaf shape formation, secondary root elongation, flowering and fertility [5]. The population of endogenous siRNAs in plants largely exceeds that of miRNAs. Increasing lines of evidence have shown that siRNAs are also important regulators in plants. The RNA silencing phenomenon was first discovered in plants as a mechanism to silence invading nucleic acids such as transgenes and viruses through the action of siRNAs [6]. It was later recognized that RNA silencing also plays important roles in the regulation of gene expression in a wide range of biological processes [7].

In recent years, extensive studies focusing on plant sRNAs have been carried out, especially in the model plant *Arabidopsis thaliana*. These studies led to the identification of a few genes related to sRNA biogenesis and functions [5,8,9], and also new groups of functional sRNAs [10-17]. For miRNAs, besides the traditionally-recognized mature miRNAs, high expression and target-cleavage function were also found for the isomiRNAs, which are sRNAs produced from the same genomic loci as miRNAs but with variable lengths, and miRNA*s, which are pairing sequences of mature miRNAs within the same hairpin-shaped precursor [18-21]. On the other hand, the known functional siRNA classes include (1) *trans*-acting siRNAs (ta-siRNAs), which are phased sRNAs produced from the *TAS* genes with the involvement of miRNAs and RNA-dependent RNA polymerase 6 (RDR6) [22], (2) repeat-associated siRNAs (ra-siRNAs), which are generated from transposons, heterochromatic and repetitive genomic regions [23], (3) natural antisense siRNAs (nat-siRNAs), which are derived from the overlapping regions of natural
antisense transcript pairs [24,25], (4) cis-acting siRNAs (ca-siRNAs) produced by the cis-cleavage of TAS3 transcript [14] and (5) DNA double-strand break induced small RNAs (diRNAs) in response to DNA damage [16]. Besides the above mentioned groups, there are also a large number of unclassified sRNAs with unknown functions.

With the application of high-throughput sequencing technology, the population and expression of sRNAs have been extensively explored in many plant species, revealing some genomic regions enriched in sRNAs [23]. These genomic regions tend to produce multiple, sometimes hundreds or thousands of, sRNAs with variable lengths. The cluster-distributed sRNAs can be grouped into small RNA clusters (SRCs), which have caught some attention in recent years [26-28]. Studies have shown that 24-nucleotides (nt) SRCs may contribute to hybrid vigor in *Arabidopsis* and maize [29,30]. Moreover, bioinformatics tools identifying SRCs have also been developed [18,31]. Here, we summarize the current knowledge and features of plant SRCs, as well as the facing challenges, with the hope to facilitate functional studies of SRCs in plants.

**Biogenesis of plant miRNAs and siRNAs**

Both miRNAs and siRNAs are derived from the double-stranded regions of RNA molecules. The difference is that miRNAs are produced from the double-stranded stem region of a single RNA molecule with hairpin-shaped precursors, whereas siRNAs are generated from double-stranded duplexes formed by two RNA molecules [5]. There are several steps involved in miRNA biogenesis. Firstly, primary miRNAs (pri-miRNAs), which are the original transcripts for miRNA processing, are transcribed in most case by RNA polymerase II (RNAPII) [32,33]. In plants, pri-miRNAs are primarily produced from intergenic regions, whereas a considerable number of animal miRNAs are encoded by introns of protein-coding genes [34]. Secondly, pri-miRNAs are processed into precursor miRNAs (pre-miRNAs) with stem-loop shaped secondary structures by the endonuclease Dicer-like 1 protein (DCL1) [35]. This process also requires the participation of many other proteins, such as the RNA-binding protein DAWDLE (DDL) [36], the C2H2-zinc finger protein SERRATE (SE) [37] and the dsRNA binding protein HYponastic LEAVES1 (HYL1) [38]. Thirdly, pre-miRNAs are processed by DCL1 into
~21-nt double-stranded miRNA duplex in nucleus [39]. The double-stranded duplex of mature miRNA and its pairing sequence, used to be named miRNA*, are then exported to the cytoplasm through the action of the HASTY protein, which is the plant ortholog of exportin 5, and other unknown factors [40]. Increasing lines of evidence have shown that the miRNA* sequences of some miRNAs are also functional under certain circumstances. The miRNA and miRNA* are now more commonly named as miRNA-5p and miRNA-3p, respectively, according to their positions on the hairpin-shaped precursor [41-43]. The produced miRNA-5p or miRNA-3p will then be methylated by HEN1 [44], a reaction that protects them from being degraded by the small RNA degrading nuclease (SDN) class of exonucleases [45]. Finally, the functional miRNA (either miRNA-5p, miRNA-3p or both) will be incorporated into Argonature (AGO) protein-centered RNA induced silencing complex (RISC) to execute their functions [46]. Classically, plant miRNAs usually associate with AGO1 to induce post-transcriptional gene silencing (PTGS) of RNA by pairing to target sequence and resulting in RNA slicing and/or translation inhibition [47,48].

On the other hand, siRNAs are generated from dsRNA duplexes and processed by one of the DCL family member proteins (Table 1). siRNA precursors include single-stranded RNA (ssRNA) molecules with near-perfect double-stranded regions formed by the fold-back of an inverted-repeat (IR) sequence [49], dsRNA duplexes formed by two RNA molecules with sequence complementarity [24] or dsRNAs converted from ssRNAs by RNA-dependent RNA polymerases (RdRPs) [50]. The resulting dsRNA molecules will be cleaved sequentially into siRNAs, and loaded to AGO proteins of the RISC complex to execute transcriptional or post-transcriptional gene silencing functions via sequencing pairing [10]. In plants, siRNAs of different lengths are produced by different DCL family proteins and exhibit different AGO protein association preferences [13]. Majority of plant siRNAs are 24-nt long, which are mainly produced by DCL3 and associate with AGO4 protein [8]. Like miRNAs, the 21-nt plant siRNAs are mainly produced by DCL1 and associate with AGO1/2 [51,52], whereas the 22-nt siRNAs are mainly produced by DCL2 and may function as backups for the 21-nt or 24-nt siRNAs [3].

**SRCs produced from miRNA precursors**
Since first being identified in 1993, the importance of miRNAs has been more and more appreciated [53,54]. With the development of high-throughput sequencing technology, many new features of miRNAs have been identified.

Mature animal miRNAs are mainly 22-nt in length; however, majority of plant miRNAs are 21-nt long and this phenomenon is more remarkable among conserved miRNAs (Figure 1A). Nearly 80% of miRNAs are 21-nt long in Arabidopsis thaliana (Ath), which is higher than that in rice and maize (Figure 1B). According to the biogenesis feature of miRNAs, high abundant sequence reads should be detected from the mature miRNA loci of miRNA precursors. For some miRNA precursors, the pairing sequence of the mature miRNA, either miRNA-5p or miRNA-3p, will also be functional therefore be protected from degradation [55]. High abundance of these mature miRNA pairing sequences has been observed under some conditions (Figure 1C). According to miRBase release 20, 32.3% mature miRNAs are reported to have their pairing miRNA-5p or miRNA-3p accumulated. Thus, the presence of both miRNA-5p and miRNA-3p sequences is also considered as strong evidence for the identification of new miRNAs [19].

Besides the functional miRNA-5p and miRNA-3p, other sRNAs with variable lengths have also been identified at or around the genomic loci of miRNA-5p and miRNA-3p, due to the imprecise cleavage of DCL enzyme. These sRNAs are named isomiRNAs [18]. Therefore, two cluster-distributed isomiRNAs can be identified in a typical miRNA precursor, whereas other regions of the precursor almost have no corresponding sRNAs produced (Figure 1C). This feature is also a critical criterion to identify bona fide miRNAs [42]. Although the sequences of isomiRNAs are either shorter or longer than those of the mature miRNAs, isomiRNAs may share the same target genes with the mature miRNAs. Nonetheless in some cases, the isomiRNAs can gain new targets [56]. In addition, expression variation among different tissues or developmental stages has also been observed for some isomiRNAs [20]. Under certain conditions, the expression of certain isomiRNAs can be even higher than that of the canonical miRNAs [56,57].

**Phased ta-siRNA clusters**

ta-siRNAs are a class of 21-nt long plant-specific siRNAs produced from long dsRNAs [22,58]. Different from other siRNAs, ta-siRNAs are derived from the cleavage product of mRNAs after
being targeted by miRNAs [10]. In general, the binding of miRNAs to specific target genes (mainly the TAS family genes) leads to the cleavage of target genes at the middle of miRNA binding sites [22]. The cleavage product of mRNAs will be converted into dsRNA by RDR6 and suppressor of gene silencing 3 (SGS3) (Table 1). The resulting dsRNA will be further processed by DCL4 in a step-wise manner to produce a phased array of 21-nt siRNAs, starting from the miRNA cleavage site (Figure 2). Therefore, a series of in-phase produced siRNAs will be identified from both strands of the dsRNA duplex converted from miRNA cleavage product [22]. These in-phase produced siRNAs are termed ta-siRNAs. Like miRNAs, ta-siRNAs also function by targeting mRNAs via non-perfect sequence complementarity and induce target cleavage at the middle of ta-siRNA binding sites [7]. Within the series of ta-siRNAs produced from the same gene, usually only one ta-siRNA is functional, which has higher expression than other ta-siRNAs [59].

ta-siRNAs were first identified in Arabidopsis in 2004 [22], and have now been extensively found in many plant species, such as rice, maize and Brassica napus [41,60,61]. The sequences and targets of ta-siRNAs also tend to be conserved across species. For example, the 5’ D7 ta-siRNA produced from TAS3 gene targets ADP-ribosylation factor 3 (ARF3) in both Arabidopsis and B. napus [41].

**22-nt SRCs derived from retrotransposons**

Although how cells distinguish sRNAs with 1-nt difference in lengths is still largely unknown, several lines of evidence have shown that the length is critical for the functions of some sRNAs. For example, 22-nt long miRNAs that are derived from the asymmetric duplex region of hairpin-shaped precursors, such as ath-miR173 and ath-miR828 [43], are capable of triggering the production of RDR6-dependent secondary siRNAs from target RNAs, whereas the corresponding 21-nt form of the same miRNA failed to do so [62,63].

The proportion of 22-nt sRNAs in maize is higher than in Arabidopsis and rice [49,64-66]. Compared to euchromatic and heterochromatic regions, 22-nt siRNAs in maize are more enriched in regions encoding retrotransposons, especially the Copia and Gypsy superfamily retrotransposons [30,67]. These 22-nt siRNAs are usually distributed in clusters on the maize
genome, therefore they can be considered as 22-nt SRCs (Figure 3). Although the generation of most endogenous siRNAs is dependent on RDR2 in Arabidopsis [68], a high abundance of 22-nt SRCs have been detected in the mediator of paramutation 1 (MOP1, homolog of RDR2) mutant, suggesting that the biogenesis of these 22-nt SRCs is independent of MOP1 [69]. Systematic studies of siRNAs between two maize inbred lines (B73 and Mo17) and their hybrid revealed that some retrotransposon-derived 22-nt SRCs are differentially expressed between the two parental lines, as well as their hybrid, suggesting their potential functions in regulating heterosis [30].

24-nt SRCs and their functions

Majority of endogenous siRNAs in plants are 24-nt long, which are primarily produced from transposons and DNA repeat sequences, especially heterochromatic regions [49]. siRNAs derived from expressed pseudogenes are also mainly 24-nt long [49]. The biogenesis of 24-nt siRNAs in plants requires the participation of a plant-specific RNA polymerase IV (RNAPIV). RNAPIV binds to the specific genomic loci through an unknown mechanism to produce ssRNAs [70]. The ssRNA will then be converted to double-stranded duplex by RDR2 and further processed into 24-nt siRNAs by DCL3 [71]. In plant, 24-nt siRNAs usually associate with AGO4 (also AGO6 and AGO9 in some cases) and play roles in RNA-directed DNA methylation (RdDM) or transposon silencing [72]. Some 24-nt siRNAs, especially those derived from pseudogenes, can also modulate histone deacetylation and histone methylation [11].

Due to the extensive existence of transposons and heterochromatic regions, 24-nt SRCs are the most commonly seen SRCs in plant genomes (Figure 4). Some 24-nt SRCs can be quite large, which span long on the genome and contain thousands of siRNAs. Several groups have studied the functions of 24-nt SRCs in regulating heterosis. Groszmann et al demonstrated that the expression of 24-nt SRCs is remarkably down-regulated in Arabidopsis hybrids [29]. It has been shown that 24-nt SRCs that are derived from transposable elements (TEs) are often inherited in an additive manner in hybrids, whereas expression of 24-nt SRCs associated with protein-coding genes that function in pathogen defense, abiotic stress tolerance and secondary metabolism is often down-regulated in hybrids [73]. In maize, 24-nt SRCs have relatively low
abundance in pericentromeric regions but are highly enriched in euchromatic regions. These 24-nt SRCs tend to be produced from the hAT, CACTA and PIF/Harbinger families of DNA transposons [67]. These studies indicate that 24-nt SRCs may play roles in regulating heterosis-related changes. However, the functions of 24-nt SRCs in other physiological processes remain to be investigated.

**Conclusion and perspectives**

The application of high-throughput sequencing technology has greatly facilitated sRNA studies, and significant progresses have been achieved during the past few years. Yet few studies on plant sRNAs have been focused on SRCs. Substantial challenges related to SRC identification and functional studies in plants still remain. For example, as many siRNAs within SRCs have multiple loci on the genome, how to discriminate the siRNA expressed loci from others is a big problem. Are siRNAs within a SRC equally important in terms of functions, or there are dominant ones? A large number of SRCs have also been identified from tRNAs and rRNAs recently, whether they are functional or only degradation products remain to be addressed as well. Nevertheless, the extensive presence of SRCs in plant genomes is a good indication for their functional importance; the remaining mysteries will be uncovered sooner or later.

**Competing interests**

The authors declare to have no competing interests.

**Acknowledgements**

This work was supported by the grants from National Science Foundation of China (Grant No. 31271349 to XJW and 31371318 to MW) and grant from Chinese Academy of Sciences (Grant No. KSCX2-EW-R-0103 to XJW).
References


[38] Kurihara Y, Takashi Y, Watanabe Y. The interaction between DCL1 and HYL1 is important for efficient and precise processing of pri-miRNA in plant microRNA biogenesis. RNA 2006;12:206-12.


Figure legends

Figure 1  Length distribution of plant miRNAs and the presence of SRCs on miRNA precursors
A. Length distribution of known miRNAs in plants and animals. B. Length distribution of known miRNAs in Arabidopsis thaliana (Ath), rice and maize. C. Typical distribution of SRCs on miRNA precursors, exemplified by ath-MIR172b, the precursor of ath-miR172b. Two SRCs formed by isomiRNAs are detected. The sequences of miRBase-recorded ath-miR172b-5p and ath-miR172b-3p are shown in red. miRNA sequences used in this analysis are collected from miRBase release 20 [43]. The pairing situation of the miRNA precursor is shown in dot-parenthesis notation, with a dot represents unpaired nucleotide and a pair of open and close parenthesis represents a pair of pairing nucleotides. The minus value shown in parenthesis to the right of the bottom line indicates the MFE of the RNA structure. miRNA, microRNAs; SRC, small RNA cluster; isomiRNA, small RNAs produced from the same genomic loci as miRNAs but with variable lengths; MFE, minimum free energy.

Figure 2  SRCs formed by ta-siRNAs
Shown are ta-siRNAs produced from TAS1a (At2g27400) gene in Arabidopsis. Small RNA data are obtained from the GEO database (accession No.GSM895899) [16]. Red and blue bars represents the expression of ta-siRNAs on the plus and minus strands of TAS1a gene, respectively, scaled from 0 – 7000 reads. Cleavage of TAS1a mRNA is triggered by miR173, the resulting sequence serves as the template to form dsRNA. The green and purple sequences represent ta-siRNAs on the TAS1a gene with adjacent ta-siRNAs indicated in different colors. ta-siRNA, trans-acting small interfering RNA; GEO, Gene Expression Omnibus; dsRNA, double-stranded RNA.

Figure 3  22-nt SRCs enriched in retrotransposons in maize
Shown in the graphs are the 22-nt SRCs enriched in the doke (A) and okor (B) family retrotransposons in maize, respectively. Expression refers to redundant sRNAs produced from the shown genomic loci, with the Y-axis represents the number of sRNAs. Sequences refer to non-redundant sRNAs produced from the shown genomic loci. siRNAs derived from the plus and minus strands of the genomic region are shown in red and blue, respectively. siRNA data are extracted from [66], only reads with sequence count > 1 are included in the analysis. SRC, small RNA cluster; siRNA, small interfering RNA.

Figure 4 Different origins of 24-nt SRCs in Arabidopsis
Examples of 24-nt SRCs derived from protein-coding gene (A), heterochromatin region (B), transposon (C) and pseudogene (D). Expression refers to redundant sRNAs produced from the shown genomic loci, with the Y-axis represents the number of sRNAs. Sequences refer to non-redundant sRNAs produced from the shown genomic loci. siRNAs derived from the plus and minus strands of the genomic region are shown in red and blue, respectively. siRNA data are extracted from [16], only reads with sequence counts > 5 are included in the analysis. Heterochromatin and euchromatin regions on chromosome 4 (Chr4, panel B) are shown in light blue and red, respectively.

Tables

Table 1 Types of small RNAs and SRCs in plants
A

B

C

ath-miR172b-5p
AGGCAGCAACAUUAGAUU......GCAGCACCAIJUAAGAUUCAC......GCAGCACAUUAAGAUUCACAU......GCAGCACAUUAAGAUUCAC......GCAGCACAUUAAGAUUCAC......

ath-miR172b-3p

<table>
<thead>
<tr>
<th>Count</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>UGAGAUCUUGAUGAUGCUGC...</td>
</tr>
<tr>
<td>515</td>
<td>GAGAAUCUUGAUGAUGCUGCAU...</td>
</tr>
<tr>
<td>63</td>
<td>AGAAUCUUGAUGAUGCUGCAU...</td>
</tr>
<tr>
<td>12</td>
<td>AGAAUCUUGAUGAUGCUGCAU...</td>
</tr>
<tr>
<td>3</td>
<td>AGAAUCUUGAUGAUGCUGCAU...</td>
</tr>
<tr>
<td>55</td>
<td>AGAAUCUUGAUGAUGCUGCAU...</td>
</tr>
<tr>
<td>49</td>
<td>AGAAUCUUGAUGAUGCUGCAU...</td>
</tr>
<tr>
<td>11298</td>
<td>AGAAUCUUGAUGAUGCUGCAU...</td>
</tr>
<tr>
<td>2892</td>
<td>AGAAUCUUGAUGAUGCUGCAU...</td>
</tr>
<tr>
<td>107</td>
<td>AGAAUCUUGAUGAUGCUGCAU...</td>
</tr>
<tr>
<td>43</td>
<td>AGAAUCUUGAUGAUGCUGCAU...</td>
</tr>
<tr>
<td>27</td>
<td>AGAAUCUUGAUGAUGCUGCAU...</td>
</tr>
<tr>
<td>583</td>
<td>AGAAUCUUGAUGAUGCUGCAU...</td>
</tr>
<tr>
<td>167</td>
<td>AGAAUCUUGAUGAUGCUGCAU...</td>
</tr>
<tr>
<td>57</td>
<td>AGAAUCUUGAUGAUGCUGCAU...</td>
</tr>
<tr>
<td>(-36.40)</td>
<td>AGGCAGCAACAUUAGAUUCACAUUGAAAUUGAUAUAUACCUAAAUUAGGGUUUUGAUAUGUAUAGAGAAUCUUGAUGAUGCUGCAUCAAC</td>
</tr>
</tbody>
</table>
TAS1a mRNA

miR173
A  Chr1: 130,280,500 bp—130,283,500 bp  
RLG_doke_AC197224

B  Chr1: 264,300,750 bp—264,301,100 bp  
RIX_okor_AC205906
Table 1 Types of sRNAs and SRCs in plants

<table>
<thead>
<tr>
<th>Type</th>
<th>Length (nt)</th>
<th>Transcribed by RNAP</th>
<th>Biogenesis enzyme</th>
<th>Processed by</th>
<th>Associate with</th>
<th>Origin</th>
<th>Function</th>
<th>Refs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>miRNA</td>
<td>~21</td>
<td>II</td>
<td>NA</td>
<td>DCL1</td>
<td>AGO1/7/10</td>
<td>ssRNA with hairpin-shaped secondary structure</td>
<td>Post-transcriptional regulation</td>
<td>[10]</td>
</tr>
<tr>
<td>ta-siRNA</td>
<td>21</td>
<td>II</td>
<td>RDR6</td>
<td>DCL4</td>
<td>AGO1</td>
<td>miRNA-cleaved TAS RNA</td>
<td>Post-transcriptional regulation</td>
<td>[22]</td>
</tr>
<tr>
<td>ra-siRNA</td>
<td>24</td>
<td>IV</td>
<td>RDR2</td>
<td>DCL3</td>
<td>AGO4</td>
<td>Repeats</td>
<td>DNA methylation</td>
<td>[23]</td>
</tr>
<tr>
<td>ca-siRNA</td>
<td>24</td>
<td>IV</td>
<td>RDR2</td>
<td>DCL3</td>
<td>AGO4</td>
<td>Transposons, repeats</td>
<td>Chromatin modification</td>
<td>[11-13]</td>
</tr>
<tr>
<td>nat-siRNA</td>
<td>21</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>AGO1</td>
<td>NA</td>
<td>Cis-cleavage TAS transcripts</td>
<td>[14]</td>
</tr>
<tr>
<td>nat-siRNA</td>
<td>21</td>
<td>NA</td>
<td>DCL1</td>
<td>NA</td>
<td>NA</td>
<td>Overlap regions of natural antisense transcripts</td>
<td>Post-transcriptional regulation</td>
<td>[15,24]</td>
</tr>
<tr>
<td>diRNA</td>
<td>21</td>
<td>IV</td>
<td>RDR2/6</td>
<td>DCL2/3/4</td>
<td>AGO2</td>
<td>DNA DSBs induced in the vicinity of DSB sites</td>
<td>Chromatin modifications or recruitment of protein complexes to DSB sites</td>
<td>[16]</td>
</tr>
<tr>
<td>easiRNA</td>
<td>21</td>
<td>II</td>
<td>RDR6</td>
<td>DCL4</td>
<td>AGO1</td>
<td>Reactivated transposons triggered by miRNAs</td>
<td>Silence TEs in epigenetically-reprogrammed germline</td>
<td>[17]</td>
</tr>
<tr>
<td>hetsiRNA</td>
<td>24</td>
<td>IV</td>
<td>RDR2</td>
<td>DCL3</td>
<td>AGO4</td>
<td>Heterochromatin</td>
<td>Heterochromatin formation</td>
<td>[10]</td>
</tr>
<tr>
<td>22-nt SRC</td>
<td>22</td>
<td>NA</td>
<td>NA</td>
<td>DCL2?</td>
<td>NA</td>
<td>Retrotransposon</td>
<td>NA</td>
<td>[30]</td>
</tr>
<tr>
<td>24-nt SRC</td>
<td>24</td>
<td>IV/V</td>
<td>RDR2</td>
<td>DCL3</td>
<td>AGO4</td>
<td>Transposon</td>
<td>DNA methylation</td>
<td>[29]</td>
</tr>
<tr>
<td>Other SRC</td>
<td>18-26</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>Highly-expressed genes</td>
<td>NA</td>
<td></td>
</tr>
</tbody>
</table>

*Note: 22-nt SRCs and 24-nt SRCs indicate SRCs mainly formed by 22-nt sRNAs and by 24-nt sRNAs, respectively, whereas other SRCs refer to SRCs comprised of sRNAs of mixed lengths but the population of sRNAs within none of length categories exceeds 50% of the total population. NA, not available. sRNA, small RNA; SRC, small RNA cluster; miRNA, microRNAs; siRNA, small interfering RNAs; ta-siRNA, trans-acting siRNA; ca-siRNA, cis-acting siRNA; ra-siRNA, repeat-associated siRNA; nat-siRNA, natural anti-sense siRNA; diRNA, double-strand break (DSB)-induced sRNA; easiRNA, epigenetically-activated siRNA; hetsiRNA, heterochromatic siRNA; RNAP, RNA polymerase; RDR, RNA-dependent RNA polymerase; DCL, Dicer-like protein; AGO, Argonaute; ssRNA, single-stranded RNA; TE, transposable element*