REVIEW

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*

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 a 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.

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 the class of most extensively studied sRNAs so far, and have been implicated 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

* Corresponding author.
   E-mail: xjwang@genetics.ac.cn (Wang XJ).

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-nucleotide (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 cases 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 sequences 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 [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 *A. 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

Table 1   Types of sRNAs and SRCs in plants

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

Note: 22-nt SRCs and 24-nt SRCs indicate SRCs mainly formed by 22-nt sRNAs and 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, microRNA; siRNA, small interfering RNA; 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.

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 iso-miRNAs 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-siR-NAs 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].
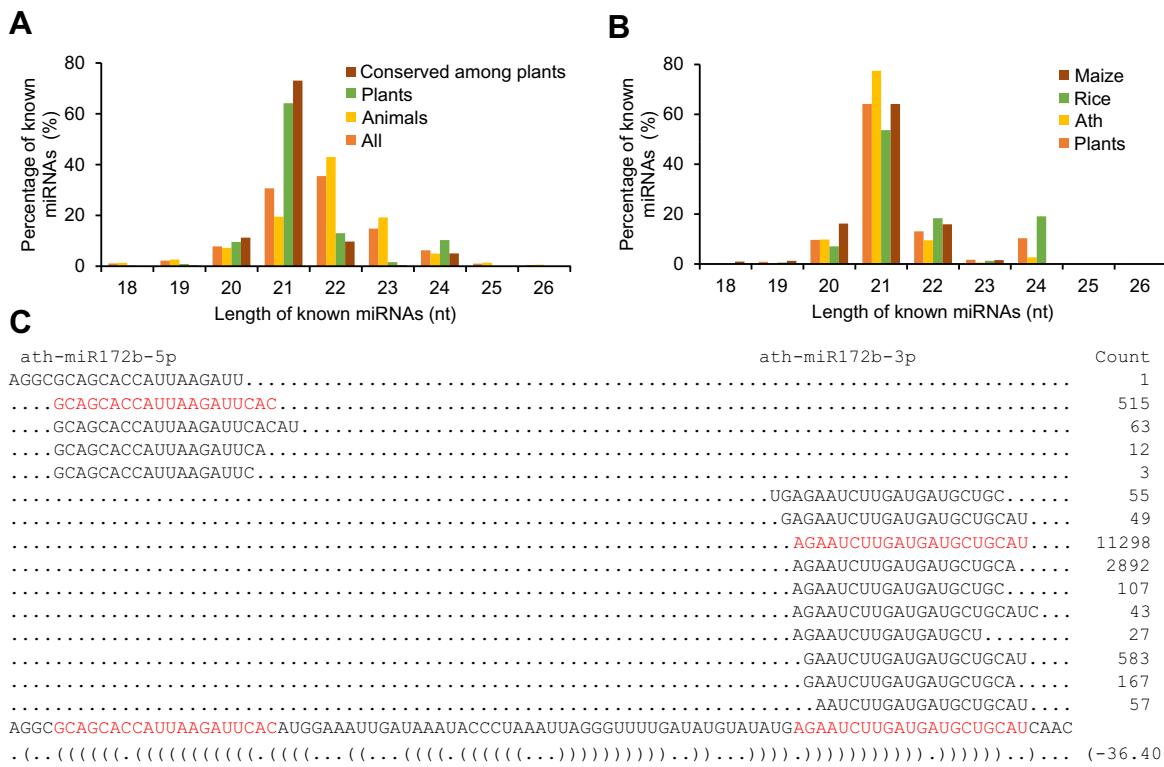
Figure 1   Length distribution of plant miRNAs and the presence of SRCs on miRNA precursors
A. Length distribution of known miRNAs in plants and animals. **B.** Length distribution of known miRNAs in *Arabidopsis thaliana* (Ath), rice and maize. **C.** Typical distribution of SRCs on miRNA precursors, 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 representing 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.
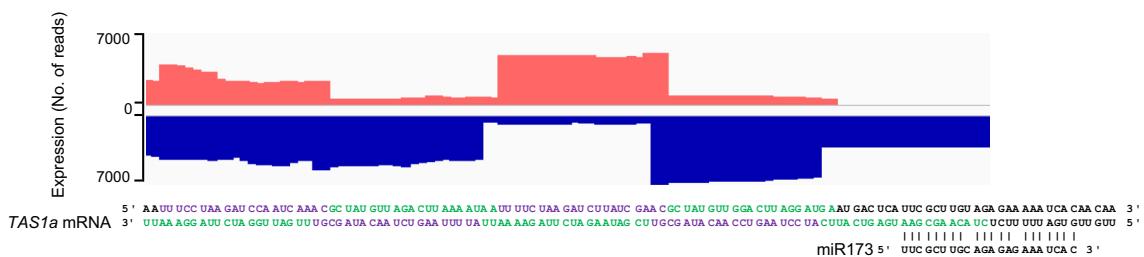
Figure 2   SRCs formed by ta-siRNAs
Shown are ta-siRNAs produced from *TAS1a* (*At2g27400*) gene in *Arabidopsis*. Small RNA data are obtained from the GEO database (accession No.GSM895899) [16]. Red and blue bars represent the expression of ta-siRNAs on the plus and minus strands of *TAS1a* gene, respectively, scaled from 0 to 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.

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 has 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].
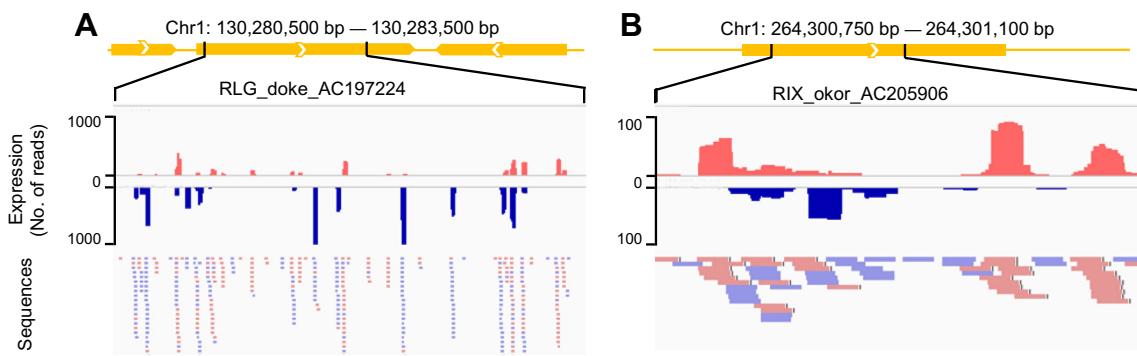
Figure 3   22-nt SRCs enriched in retrotransposons in maize
Shown in the graphs are the 22-nt SRCs enriched in the doke (**A**) and okor (**B**) family retrotransposons in maize, respectively. Expression refers to redundant sRNAs produced from the shown genomic loci, with the Y-axis representing the 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.
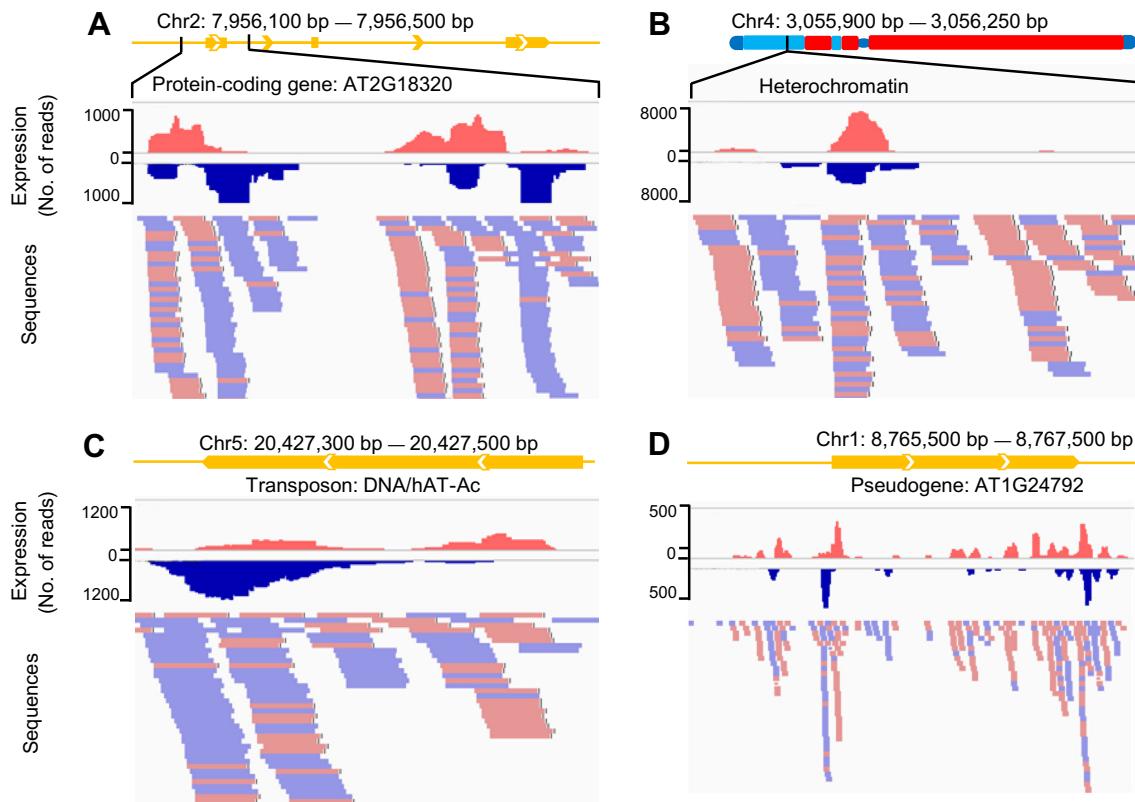
Figure 4   Different origins of 24-nt SRCs in *Arabidopsis*
Examples of 24-nt SRCs derived from protein-coding gene (**A**), heterochromatin region (**B**), transposon (**C**) and pseudogene (**D**). Expression refers to redundant sRNAs produced from the shown genomic loci, with the Y-axis representing 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.

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]. siR-NAs 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 derived from transposable elements (TEs) are often inherited in an additive manner in hybrids, whereas the 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

References

[1] Moldovan D, Spriggs A, Yang J, Pogson BJ, Dennis ES, Wilson IW. Hypoxia-responsive microRNAs and *trans*-acting small interfering RNAs in *Arabidopsis*. J Exp Bot 2010;61:165–77.

[2] Chuck G, Candela H, Hake S. Big impacts by small RNAs in plant development. Curr Opin Plant Biol 2009;12:81–6.

[3] Parent JS, Martínez de Alba AE, Vaucheret H. The origin and effect of small RNA signaling in plants. Front Plant Sci 2012;3:179.

[4] Baumberger N, Baulcombe DC. *Arabidopsis* ARGONAUTE1 is an RNA Slicer that selectively recruits microRNAs and short interfering RNAs. Proc Natl Acad Sci U S A 2005;102:11928–33.

[5] Voinnet O. Origin, biogenesis, and activity of plant microRNAs. Cell 2009;136:669–87.

[6] Vaucheret H, Béclin C, Fagard M. Post-transcriptional gene silencing in plants. J Cell Sci 2001;114:3083–91.

[7] Carthew RW, Sontheimer EJ. Origins and mechanisms of miRNAs and siRNAs. Cell 2009;136:642–55.

[8] Castel SE, Martienssen RA. RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond. Nat Rev Genet 2013;14:100–12.

[9] Liu YX, Chang W, Han YP, Zou Q, Guo MZ, Li WB. *In silico* detection of novel microRNAs genes in soybean genome. Agric Sci China 2011;10:1336–45.

[10] Chen X. Small RNAs and their roles in plant development. Annu Rev Cell Dev Biol 2009;25:21–44.

[11] Zilberman D, Cao X, Jacobsen SE. ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. Science 2003;299:716–9.

[12] Chan SW, Zilberman D, Xie Z, Johansen LK, Carrington JC, Jacobsen SE. RNA silencing genes control *de novo* DNA methylation. Science 2004;303:1336.

[13] Ghildiyal M, Zamore PD. Small silencing RNAs: an expanding universe. Nat Rev Genet 2009;10:94–108.

[14] Zhang C, Li G, Wang J, Zhu S, Li H. Cascading *cis*-cleavage on transcript from *trans*-acting siRNA-producing locus 3. Int J Mol Sci 2013;14:14689–99.

[15] Ron M, Alandete Saez M, Eshed Williams L, Fletcher JC, McCormick S. Proper regulation of a sperm-specific *cis*-nat-siRNA is essential for double fertilization in *Arabidopsis*. Genes Dev 2010;24:1010–21.

[16] Wei W, Ba Z, Gao M, Wu Y, Ma Y, Amiard S, et al. A role for small RNAs in DNA double-strand break repair. Cell 2012;149:101–12.

[17] Creasey KM, Zhai J, Borges F, Van Ex F, Regulski M, Meyers BC, et al. MiRNAs trigger widespread epigenetically activated siRNAs from transposons in *Arabidopsis*. Nature 2014;508:411–5.

[18] Luo GZ, Yang W, Ma YK, Wang XJ. ISRNA: an integrative online toolkit for short reads from high-throughput sequencing data. Bioinformatics 2014;30:434–6.

[19] Meyers BC, Axtell MJ, Bartel B, Bartel DP, Baulcombe D, Bowman JL, et al. Criteria for annotation of plant microRNAs. Plant Cell 2008;20:3186–90.

[20] Lelandais-Brière C, Naya L, Sallet E, Calenge F, Frugier F, Hartmann C, et al. Genome-wide medicago truncatula small RNA analysis revealed novel microRNAs and isoforms differentially regulated in roots and nodules. Plant Cell 2009;21:2780–96.

[21] Guo L, Lu Z. Global expression analysis of miRNA gene cluster and family based on isomiRs from deep sequencing data. Comput Biol Chem 2010;34:165–71.

[22] Allen E, Xie Z, Gustafson AM, Carrington JC. MicroRNA-directed phasing during *trans*-acting siRNA biogenesis in plants. Cell 2005;121:207–21.

[23] Vazquez F. *Arabidopsis* endogenous small RNAs: highways and byways. Trends Plant Sci 2006;11:460–8.

[24] Borsani O, Zhu J, Verslues PE, Sunkar R, Zhu JK. Endogenous siRNAs derived from a pair of natural *cis*-antisense transcripts regulate salt tolerance in *Arabidopsis*. Cell 2005;123:1279–91.

[25] Katiyar-Agarwal S, Morgan R, Dahlbeck D, Borsani O, Villegas Jr A, Zhu JK, et al. A pathogen-inducible endogenous siRNA in plant immunity. Proc Natl Acad Sci U S A 2006;103:18002–7.

[26] He G, Zhu X, Elling AA, Chen L, Wang X, Guo L, et al. Global epigenetic and transcriptional trends among two rice subspecies and their reciprocal hybrids. Plant Cell 2010;22:17–33.

[27] Liu YX, Han YP, Chang W, Zou Q, Guo MZ, Li WB. Genomic analysis of microRNA promoters and their *cis*-acting elements in soybean. Agric Sci China 2010;9:1561–70.

[28] Wang H, Zhang X, Liu J, Kiba T, Woo J, Ojo T, et al. Deep sequencing of small RNAs specifically associated with *Arabidopsis* AGO1 and AGO4 uncovers new AGO functions. Plant J 2011;67:292–304.

[29] Groszmann M, Greaves IK, Albertyn ZI, Scofield GN, Peacock WJ, Dennis ES. Changes in 24-nt siRNA levels in *Arabidopsis* hybrids suggest an epigenetic contribution to hybrid vigor. Proc Natl Acad Sci U S A 2011;108:2617–22.

[30] Barber WT, Zhang W, Win H, Varala KK, Dorweiler JE, Hudson ME, et al. Repeat associated small RNAs vary among parents and following hybridization in maize. Proc Natl Acad Sci U S A 2012;109:10444–9.

[31] Wu HJ, Ma YK, Chen T, Wang M, Wang XJ. PsRobot: a web-based plant small RNA meta-analysis toolbox. Nucleic Acids Res 2012;40:W22–8.

[32] Lee Y, Kim M, Han J, Yeom KH, Lee S, Baek SH, et al. MicroRNA genes are transcribed by RNA polymerase II. EMBO J 2004;23:4051–60.

[33] Li J, Liu YX, Han YP, Li YG, Guo MZ, Li WB. MicroRNA primary transcripts and promoter elements analysis in soybean (*Glycine max* L. Merril.). J Integr Agric 2013;12:1522–9.

[34] Ruby JG, Jan CH, Bartel DP. Intronic microRNA precursors that bypass Drosha processing. Nature 2007;448:83–6.

[35] Kurihara Y, Watanabe Y. *Arabidopsis* micro-RNA biogenesis through Dicer-like 1 protein functions. Proc Natl Acad Sci U S A 2004;101:12753–8.

[36] Yu B, Bi L, Zheng B, Ji L, Chevalier D, Agarwal M, et al. The FHA domain proteins DAWDLE in *Arabidopsis* and SNIP1 in humans act in small RNA biogenesis. Proc Natl Acad Sci U S A 2008;105:10073–8.

[37] Lobbes D, Rallapalli G, Schmidt DD, Martin C, Clarke J. SERRATE: a new player on the plant microRNA scene. EMBO Rep 2006;7:1052–8.

[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.

[39] Bartel DP. MicroRNAs, genomics, biogenesis, mechanism, and function. Cell 2004;116:281–97.

[40] Bollman KM, Aukerman MJ, Park MY, Hunter C, Berardini TZ, Poethig RS. HASTY, the *Arabidopsis* ortholog of exportin 5/MSN5, regulates phase change and morphogenesis. Development 2003;130:1493–504.

[41] Zhao YT, Wang M, Fu SX, Yang WC, Qi CK, Wang XJ. Small RNA profiling in two *Brassica napus* cultivars identifies microRNAs with oil production- and development-correlated expression and new small RNA classes. Plant Physiol 2012;158:813–23.

[42] Kozomara A, Griffiths-Jones S. MiRBase: annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res 2014;42:D68–73.

[43] Kozomara A, Griffiths-Jones S. MiRBase: integrating microRNA annotation and deep-sequencing data. Nucleic Acids Res 2011;39:D152–7.

[44] Yu B, Bi L, Zhai J, Agarwal M, Li S, Wu Q, et al. SiRNAs compete with miRNAs for methylation by HEN1 in *Arabidopsis*. Nucleic Acids Res 2010;38:5844–50.

[45] Ramachandran V, Chen X. Degradation of microRNAs by a family of exoribonucleases in *Arabidopsis*. Science 2008;321:1490–2.

[46] Jones-Rhoades MW, Bartel DP, Bartel B. MicroRNAs and their regulatory roles in plants. Annu Rev Plant Biol 2006;57:19–53.

[47] Brodersen P, Sakvarelidze-Achard L, Bruun-Rasmussen M, Dunoyer P, Yamamoto YY, Sieburth L, et al. Widespread translational inhibition by plant miRNAs and siRNAs. Science 2008;320:1185–90.

[48] Li S, Liu L, Zhuang X, Yu Y, Liu X, Cui X, et al. MicroRNAs inhibit the translation of target mRNAs on the endoplasmic reticulum in *Arabidopsis*. Cell 2013;153:562–74.

[49] Kasschau KD, Fahlgren N, Chapman EJ, Sullivan CM, Cumbie JS, Givan SA, et al. Genome-wide profiling and analysis of *Arabidopsis* siRNAs. PLoS Biol 2007;5:e57.

[50] Tabach Y, Billi AC, Hayes GD, Newman MA, Zuk O, Gabel H, et al. Identification of small RNA pathway genes using patterns of phylogenetic conservation and divergence. Nature 2013;493:694–8.

[51] Qi Y, Denli AM, Hannon GJ. Biochemical specialization within *Arabidopsis* RNA silencing pathways. Mol Cell 2005;19:421–8.

[52] Mi SJ, Cai T, Hu Y, Chen Y, Hodges E, Ni F, et al. Sorting of small RNAs into *Arabidopsis* argonaute complexes is directed by the 5′ terminal nucleotide. Cell 2008;133:116–27.

[53] Lee RC, Feinbaum RL, Ambros V, The C. *Elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. Cell 1993;75:843–54.

[54] Reinhart BJ, Slack FJ, Basson M, Pasquinelli AE, Bettinger JC, Rougvie AE, et al. The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. Nature 2000;403:901–6.

[55] Yu B, Yang Z, Li J, Minakhina S, Yang M, Padgett RW, et al. Methylation as a crucial step in plant microRNA biogenesis. Science 2005;307:932–5.

[56] Galli V, Guzman F, de Oliveira LF, Loss-Morais G, Körbes AP, Silva SD, et al. Identifying microRNAs and transcript targets in *Jatropha* seeds. PLoS One 2014;9:e83727.

[57] Zhang W, Gao S, Zhou X, Xia J, Chellappan P, Zhou X, et al. Multiple distinct small RNAs originate from the same microRNA precursors. Genome Biol 2010;11:R81.

[58] Peragine A, Yoshikawa M, Wu G, Albrecht HL, Poethig RS. SGS3 and SGS2/SDE1/RDR6 are required for juvenile development and the production of *trans*-acting siRNAs in *Arabidopsis*. Genes Dev 2004;18:2368–79.

[59] Axtell MJ, Jan C, Rajagopalan R, Bartel DP. A two-hit trigger for siRNA biogenesis in plants. Cell 2006;127:565–77.

[60] Nagasaki H, Itoh J, Hayashi K, Hibara K, Satoh-Nagasawa N, Nosaka M, et al. The small interfering RNA production pathway is required for shoot meristem initiation in rice. Proc Natl Acad Sci U S A 2007;104:14867–71.

[61] Nogueira FT, Chitwood DH, Madi S, Ohtsu K, Schnable PS, Scanlon MJ, et al. Regulation of small RNA accumulation in the maize shoot apex. PLoS Genet 2009;5:e1000320.

[62] Cuperus JT, Carbonell A, Fahlgren N, Garcia-Ruiz H, Burke RT, Takeda A, et al. Unique functionality of 22-nt miRNAs in triggering RDR6-dependent siRNA biogenesis from target transcripts in *Arabidopsis*. Nat Struct Mol Biol 2010;17:997–1003.

[63] Chen HM, Chen LT, Patel K, Li YH, Baulcombe DC, Wu SH. 22-Nucleotide RNAs trigger secondary siRNA biogenesis in plants. Proc Natl Acad Sci U S A 2010;107:15269–74.

[64] Ding D, Wang Y, Han M, Fu Z, Li W, Liu Z, et al. MicroRNA transcriptomic analysis of heterosis during maize seed germination. PLoS One 2012;7:e39578.

[65] Chen CJ, Liu Q, Zhang YC, Qu LH, Chen YQ, Gautheret D. Genome-wide discovery and analysis of microRNAs and other small RNAs from rice embryogenic callus. RNA Biol 2011;8:538–47.

[66] Wang L, Liu H, Li D, Chen H. Identification and characterization of maize microRNAs involved in the very early stage of seed germination. BMC Genomics 2011;12:154.

[67] He G, Chen B, Wang X, Li X, Li J, He H, et al. Conservation and divergence of transcriptomic and epigenomic variation in maize hybrids. Genome Biol 2013;14:R57.

[68] Jia Y, Lisch DR, Ohtsu K, Scanlon MJ, Nettleton D, Schnable PS. Loss of RNA-dependent RNA polymerase 2 (RDR2) function causes widespread and unexpected changes in the expression of transposons, genes, and 24-nt small RNAs. PLoS Genet 2009;5: e1000737.

[69] Nobuta K, Lu C, Shrivastava R, Pillay M, De Paoli E, Accerbi M, et al. Distinct size distribution of endogenous siRNAs in maize: evidence from deep sequencing in the mop1-1 mutant. Proc Natl Acad Sci U S A 2008;105:14958–63.

[70] Mosher RA, Schwach F, Studholme D, Baulcombe DC. PolIVb influences RNA-directed DNA methylation independently of its role in siRNA biogenesis. Proc Natl Acad Sci U S A 2008;105: 3145–50.

[71] Nagano H, Fukudome A, Hiraguri A, Moriyama H, Fukuhara T. Distinct substrate specificities of *Arabidopsis* DCL3 and DCL4. Nucleic Acids Res 2014;42:1845–56.

[72] Qi Y, He X, Wang XJ, Kohany O, Jurka J, Hannon GJ. Distinct catalytic and non-catalytic roles of ARGONAUTE4 in RNA-directed DNA methylation. Nature 2006;443: 1008–12.

[73] Li Y, Varala K, Moose SP, Hudson ME. The inheritance pattern of 24 nt siRNA clusters in *Arabidopsis* hybrids is influenced by proximity to transposable elements. PLoS One 2012;7:e47043.