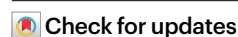


Parallel degradome-seq and DMS-MaPseq substantially revise the miRNA biogenesis atlas in *Arabidopsis*

Received: 16 January 2024

Accepted: 16 May 2024

Published online: 25 June 2024



Xingxing Yan^{1,2,11}, Changhao Li^{1,11}, Kaiye Liu^{1,8}, Tianru Zhang^{1,2}, Qian Xu^{1,3}, Xindi Li¹, Jiaying Zhu¹, Ziyang Wang^{4,9}, Anikah Yusuf^{1,10}, Shuqing Cao⁵, Xu Peng⁶, James J. Cai^{3,7} & Xiuren Zhang^{1,2,4}✉

MicroRNAs (miRNAs) are produced from highly structured primary transcripts (pri-miRNAs) and regulate numerous biological processes in eukaryotes. Due to the extreme heterogeneity of these structures, the initial processing sites of plant pri-miRNAs and the structural rules that determine their processing have been predicted for many miRNAs but remain elusive for others. Here we used semi-active DCL1 mutants and advanced degradome-sequencing strategies to accurately identify the initial processing sites for 147 of 326 previously annotated *Arabidopsis* miRNAs and to illustrate their associated pri-miRNA cleavage patterns. Elucidating the in vivo RNA secondary structures of 73 pri-miRNAs revealed that about 95% of them differ from in silico predictions, and that the revised structures offer clearer interpretation of the processing sites and patterns. Finally, DCL1 partners Serrate and HYL1 could synergistically and independently impact processing patterns and in vivo RNA secondary structures of pri-miRNAs. Together, our work sheds light on the precise processing mechanisms of plant pri-miRNAs.

MicroRNAs (miRNAs) regulate numerous biological processes in plants and animals^{1–3}. Mature miRNAs are loaded into Argonaute (AGO) protein to form RNA-induced silencing complex (RISC), which fulfils slicing or translational repression of the target transcripts on the basis of sequence complementarity^{4,5}. The production of functional miRNAs entails precise and efficient processing of primary miRNAs (pri-miRNAs) by Microprocessor. Pri-miRNAs are characterized by stem–loop structures consisting of flanking single-stranded (ss) basal segments, a lower stem, a double-stranded (ds) duplex of miRNA and

its complementary strand (miRNA/*), an upper stem, and a terminal loop⁶. In metazoans, the structures of pri-miRNAs are relatively uniform^{7–11}. By contrast, plant pri-miRNAs display remarkable diversity in shapes and sizes, with variable positioning of miRNA/* duplexes, even within the same *MIRNA* families⁶. Due to structural heterogeneity, plant pri-miRNAs can be processed from base-to-loop (BTL) and sequential BTL (SBTL) or from loop-to-base (LTB) and sequential LTB (SLTB)^{12–15}. Notably, terminal-loop-branched pri-miRNAs can be processed bidirectionally from either BTL or LTB, frequently resulting in productive

¹Department of Biochemistry and Biophysics, Texas A&M University, College Station, TX, USA. ²Molecular and Environmental Plant Sciences, Texas A&M University, College Station, TX, USA. ³Department of Veterinary Integrative Biosciences, Texas A&M University, College Station, TX, USA. ⁴Department of Biology, Texas A&M University, College Station, TX, USA. ⁵School of Food and Biological Engineering, Hefei University of Technology, Hefei, China. ⁶Department of Medical Physiology, School of Medicine, Texas A&M University, Bryan, TX, USA. ⁷Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX, USA. ⁸Present address: National Key Laboratory for Tropical Crop Breeding, School of Tropical Agriculture and Forestry, Hainan University, Haikou, China. ⁹Present address: McWilliams School of Biomedical Informatics, University of Texas Health Science Center at Houston, Houston, TX, USA. ¹⁰Present address: Doctor of Osteopathic Medicine Program, Des Moines University, West Des Moines, IA, USA. ¹¹These authors contributed equally: Xingxing Yan, Changhao Li. ✉e-mail: xiuren.zhang@tamu.edu

and abortive processing of miRNAs, respectively⁶. Plant Microprocessor canonically cuts pri-miRNAs at a site 15–17 nucleotides (nt) from a reference ssRNA–dsRNA junction region^{6,16} (Fig. 1a). Paradoxically, numerous pri-miRNAs lack the reference region within the expected distance according to in silico RNA modelling. Also, RNA secondary structures (RSS) of pri-miRNAs are dynamic and can be remodelled by RNA helicases, as exemplified by CHR2 in vivo¹⁷. How these pri-miRNAs are recognized and precisely processed by Microprocessor in vivo remains elusive. In addition, due to the RSS complexity of pri-miRNAs and challenges in identifying the processing patterns, the bona fide identities of numerous annotated miRNAs are still inconclusive^{18–20}. Furthermore, many pri-miRNAs contain extended hairpin structures, and their first cutting sites and the reference structures determining the initial cleavages remain to be determined.

Arabidopsis Microprocessor comprises Dicer-like1 (DCL1) and two core cofactors, zinc-finger protein Serrate (SE) and a dsRNA-binding protein called hyponastic leaves1 (HYL1)^{21–23}. Microprocessor sequentially processes pri-miRNAs to precursors of miRNAs (pre-miRNAs) and finally to miRNA/* duplexes²⁴. DCL1 harbours two RNase III domains—the RIII a and RIII b domains. We have previously identified two key residues, Glu1507 and Glu1696, located in RIII a and b, respectively, as critical to the enzymatic activity of this protein⁶. The substitution of glutamine for glutamate at either of the residues leads to two semi-active DCL1 point mutants (E1507Q and E1696Q, respectively) that only cut one strand of pri-miRNA duplexes and fail to proceed to further processing, resulting in partially processed pri-miRNA intermediates. This abortive processing event would create an opportunity to pinpoint the first cleavage sites by Microprocessor in vivo. Both HYL1 and SE have been reported to stimulate cleavage of the precursors by DCL1 and to improve its accuracy^{25–27}. This notwithstanding, their contributions to the molecular ruler of plant Microprocessor and impacts on RSS of pri-miRNAs are poorly understood.

Here we pinpointed the first cleavage sites of pri-miRNAs via degradome sequencing (degradome-seq) of lines expressing semi-active DCL1 variants. We clarified the 147 bona fide substrates of DCL1 from the 326 annotated pri-miRNAs and then comprehensively re-sorted the processing patterns of the true pri-miRNAs. In parallel, we conducted DMS–MaPseq²⁸ to decipher the in vivo RSS of pri-miRNAs. Notably, 69 of 73 detectable pri-miRNAs (~95%) displayed RSS that deviated to varying degrees from the in silico predicted RSS, providing a better interpretation of the initial cleavage sites for numerous pri-miRNAs. Surprisingly, whereas numerous pri-miRNAs contain the canonical reference sites to direct Microprocessor for cleavage, they tend to harbour internal loops or bulges that are 9–11 nt away from the initial cleavage sites and might be meaningful in plants. Furthermore, approximately 77% of pri-miRNAs underwent their first cuttings at internal loops or bulge regions, providing a new guideline for artificial miRNA design. We also found that SE and HYL1 imposed varying impacts on the effectiveness and precision as well as in vivo RSS of different pri-miRNAs.

This comprehensive study thus substantially revises the current model of pri-miRNA processing from different perspectives in plants.

Results

Degradome-seq defines the first cutting sites of pri-miRNAs

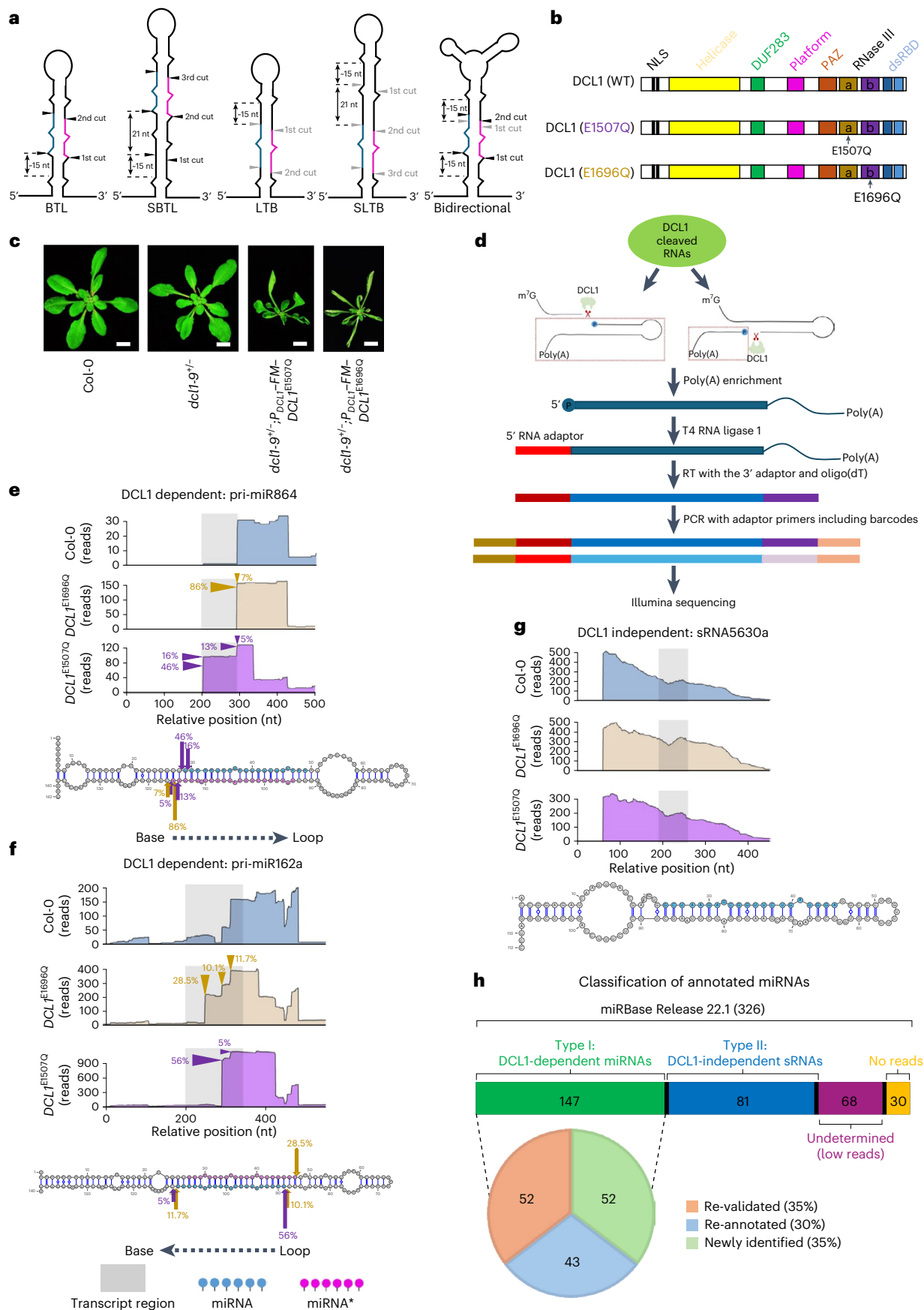
To systematically pinpoint the first cleavage sites of pri-miRNAs, we transformed native-promoter-driven Flag–4Myc (FM) tagged semi-active forms of DCL1-E1507Q and DCL1-E1696Q into *dcl1-9* heterozygotes (Fig. 1b and Supplementary Fig. 1a,b). Since *dcl1-9* harbours a transfer DNA insertion in the second dsRNA-binding domain (exon 19) and its homozygotes are embryonic lethal, we initially assumed that the introduction of *DCL1*^{E1507Q} and *DCL1*^{E1696Q} into *dcl1-9*^{+/–} would generate hypomorphic alleles. Among hundreds of complementation lines, we were able to obtain numerous transgenic lines with clear developmental defects mimicking *dcl1-9*^{+/–} lines. However, all these plants were genotyped as *dcl1-9*^{+/–}; *DCL1*^{E1507Q} and *dcl1-9*^{+/–}; *DCL1*^{E1696Q} lines, which still harboured a copy of the wild-type (WT) *DCL1* allele (Fig. 1c). Furthermore, the severity of the plant growth abnormality was correlated with the expression level of the semi-active DCL1 variants and reduced levels of miRNAs in the transgenic lines (Supplementary Fig. 1a,b). The semi-active variants of DCL1 thus served as dominant-negative forms to compete with the WT DCL1 protein for pri-miRNAs, leading to the accumulation of a series of intermediately processed forms of pri-miRNAs with descending orders of amounts in the transgenic lines (Supplementary Fig. 1c). From these materials, we recovered 3' ends of cleaved pri-miRNA segments, performed degradome-seq, and obtained in-depth and high-quality datasets (Methods) (Fig. 1d and Supplementary Fig. 1d,e).

Bona fide miRNAs based on the cleavage activity of DCL1

We extended the reads 200 bp upstream and downstream of the annotated pri-miRNAs, aiming at capturing the possible cleavage sites in a broader context flanking the pri-miRNAs. Mapping of sequence reads onto the extended regions revealed that the patterns of reads' distributions are remarkably diversified among different pri-miRNAs (Fig. 1e–g). For two canonical pri-miRNAs, pri-miR864 and pri-miR162a, we were able to detect some remnants of DCL1-processed intermediates in WT plants (Fig. 1e,f, top). By contrast, we clearly observed several abrupt steps of read accumulation in pri-miRNA regions (the purple and yellow-brown arrows in the shaded transcript areas) in the complementation lines expressing DCL1-E1507Q and DCL1-E1696Q, respectively (Fig. 1e,f, middle two panels). The abrupt increase of reads indicates that intermediate pri-miRNA products, resulting from a series of abortive processing steps by the DCL1 complex, were substantially accumulated in DCL1 semi-active mutants, enabling us to track sequential activities of DCL1 on pri-miRNAs. The steep patterns of read accumulations along pri-miRNAs could also serve as a benchmark to identify bona fide substrates of DCL1 and resultant miRNAs in vivo. With this criterion, we surveyed all 326 previously annotated pri-miRNAs from the

Fig. 1 | Identification of bona fide pri-miRNAs by degradome-seq of semi-active DCL1 mutants in *Arabidopsis*. **a**, Five processing patterns of pri-miRNAs: BTL, SBTL, LTB, SLTB and bidirectional processing. In each case, DCL1 initially cleaves pri-miRNAs at a site that is ~15–17 nt away from a reference ssRNA–dsRNA junction (15–17-nt molecular ruler). The blue and pink regions in the pri-miRNA diagrams represent miRNA/* duplexes. Black arrowheads and text in the cartoon indicate the cleavage direction from base to loop, while grey arrowheads and text represent the cleavage pattern from loop to base. **b**, The domain arrangement of DCL1 (WT) and point alterations in semi-active DCL1 mutants (E1507Q and E1696Q). The nuclear localization signal (NLS), helicase domain, DUF283, Platform, PAZ, RNase III a and III b, and dsRNA-binding domain (dsRBD) are shown in black, yellow, green, pink, dark orange, yellow-brown, purple and blue, respectively. **c**, Phenotypes of three-week-old Col-0, *dcl1-9*^{+/–} and transgenic *dcl1-9*^{+/–}; *P*_{DCL1}–FM–*DCL1*^{E1507Q} and *dcl1-9*^{+/–}; *P*_{DCL1}–FM–*DCL1*^{E1696Q} plants. Scale bars, 1 cm. **d**, Scheme for construction of the degradome-seq libraries. RT,

reverse transcription. **e–g**, Exemplified cleavage patterns of DCL1-dependent pri-miR864 (BTL direction) (**e**), DCL1-dependent pri-miR162a (LTB direction) (**f**) and DCL1-independent sRNA5630a (**g**). The yellow-brown and purple arrows represent RNase III a and III b cleavage sites, respectively. The percentages indicate the relative cutting ratios. Note that the previously claimed pri-miR5630a does not have a clear cleavage site by DCL1 in the WT or the *DCL1*^{E1507Q} and *DCL1*^{E1696Q} lines. **h**, The fraction of pri-miRNAs that have been verified to be DCL1 dependent. The total number of *Arabidopsis* pri-miRNAs is taken from miRBase Release 22.1. Note that only 147 of the 326 previously annotated pri-miRNAs are bona fide miRNAs on the basis of the cleavage activity of DCL1. The processing patterns for 35% of the true pri-miRNAs could be re-validated, whereas the other 30% display processing patterns deviating from the published literature and have now been re-annotated. Additionally, the processing patterns of the other 35% of pri-miRNAs were newly identified here.



miRBase Release 22.1 and found that a total of 147 pri-miRNAs displayed clear processing patterns in the two DCL1 semi-active mutants. These pri-miRNAs were thus confirmed to be genuine targets of DCL1, and their products are bona fide miRNAs (Type I: DCL1-dependent miRNAs) (Fig. 1h and Supplementary Table 1).

Intriguingly, we found that 81 of the 326 annotated pri-miRNAs, exemplified by pri-miR5630a, did not show any DCL1 cleavage activity in the DCL1 semi-active mutants and WT plants, as the read accumulations are gradually distributed along the reference sequence (Fig. 1g,h). This group of small RNAs (sRNAs) should thus be re-annotated to DCL1-independent sRNAs. In addition, 98 pri-miRNAs remained to be re-annotated because the read counts of the processed pri-miRNAs were extremely low (less than 15 reads) or barely detectable in this study.

Algorithm for classifying the processing patterns of pri-miRNAs

We next pinpointed the first cleavage positions for 147 true pri-miRNAs. We calculated the ratios of cleavages at the sites of individual nucleotides by dividing the accumulated reads at each position by the total accumulated reads across the extended pri-miRNAs (Supplementary Fig. 1f). Because the dominant-negative form of DCL1 yielded the largest amount of abortive processing products after the first cleavage that could not proceed to the next round of processing, the nucleotides with the highest cutting ratios should be defined as the first cleavage sites along the pri-miRNAs. For instance, pri-miR864 displayed a predominant cutting (86%) at one position and a satellite cutting at a nearby site (7%) along the extended pri-miRNAs in the *DCL1^{E1696Q}* transgenic line. This position should be counted as the first cleavage site on one strand of the pri-miR864 duplex. In parallel, a predominant cutting site (46%) and its neighbourhood with a 16% cutting ratio were detected in the *DCL1^{E1507Q}* line, and this position should be the first cut on the other strand of pri-miR864. Two minor cutting events (13% and 5%) were also detected in the *DCL1^{E1507Q}* line, clearly resulting from the activity of residual WT DCL1 protein in the dominant-negative line (Fig. 1e, bottom). Concomitant alignment of the two predominant cutting sites on the two strands of the pri-miR864 duplex placed the first cleavage sites close to the base region relative to the miR864/* duplex. This type of processing pattern indicates that Microprocessor processes pri-miR864 from a BTL direction, as reported previously^{12,14,15}.

Pri-miR162a displayed the highest cutting frequency of 56% at one site in the *DCL1^{E1507Q}* line, representing the first cleavage on one strand of the pri-miR162a duplex. The second-largest cleavage event (5%) took place 21 nt downstream due to the activity of the WT DCL1 protein in the line. Similarly, we detected the highest cutting frequency of 28.5%, indicative of the first processing site on the other strand of the pri-miR162a duplex, in the *DCL1^{E1696Q}* line. Again, two additional cutting sites with ratios of approximately 10.1% and 11.7% with a 21-nt internal distance on one strand of the duplex were observed due to one copy of

WT DCL1 protein in the line, also because only the poly(A)-harboured debris of pri-miRNAs processed from Microprocessor could be recovered in the assay (Fig. 1f, bottom). Since the two cleavage sites with the highest cutting frequency were proximal to the top region relative to the miR162/* duplex, and the secondary cleavage sites were distal from the top region, this type of processing is considered to be a standardized LTB pattern, as observed in previous studies^{14,15}.

We also found that DCL1 RNase domain b first cleaved at the 5' arm of pri-miRNA in the *DCL1^{E1507Q}* line, while DCL1 RNase domain a fulfilled the first cleavage at the 3' arm in the *DCL1^{E1696Q}* line for the BTL-type pri-miRNAs. By contrast, the DCL1 RNase domains a and b cut the 5' and 3' arms, respectively, for the LTB-type pri-miRNAs. This finding aligns with the in vitro conclusion from our previous work⁶. Taken together, our degradome analysis of pri-miR864 and pri-miR162a served as a benchmark for annotating the processing patterns and exploring new features of all 147 bona fide pri-miRNAs.

The processing atlas of 147 pri-miRNAs

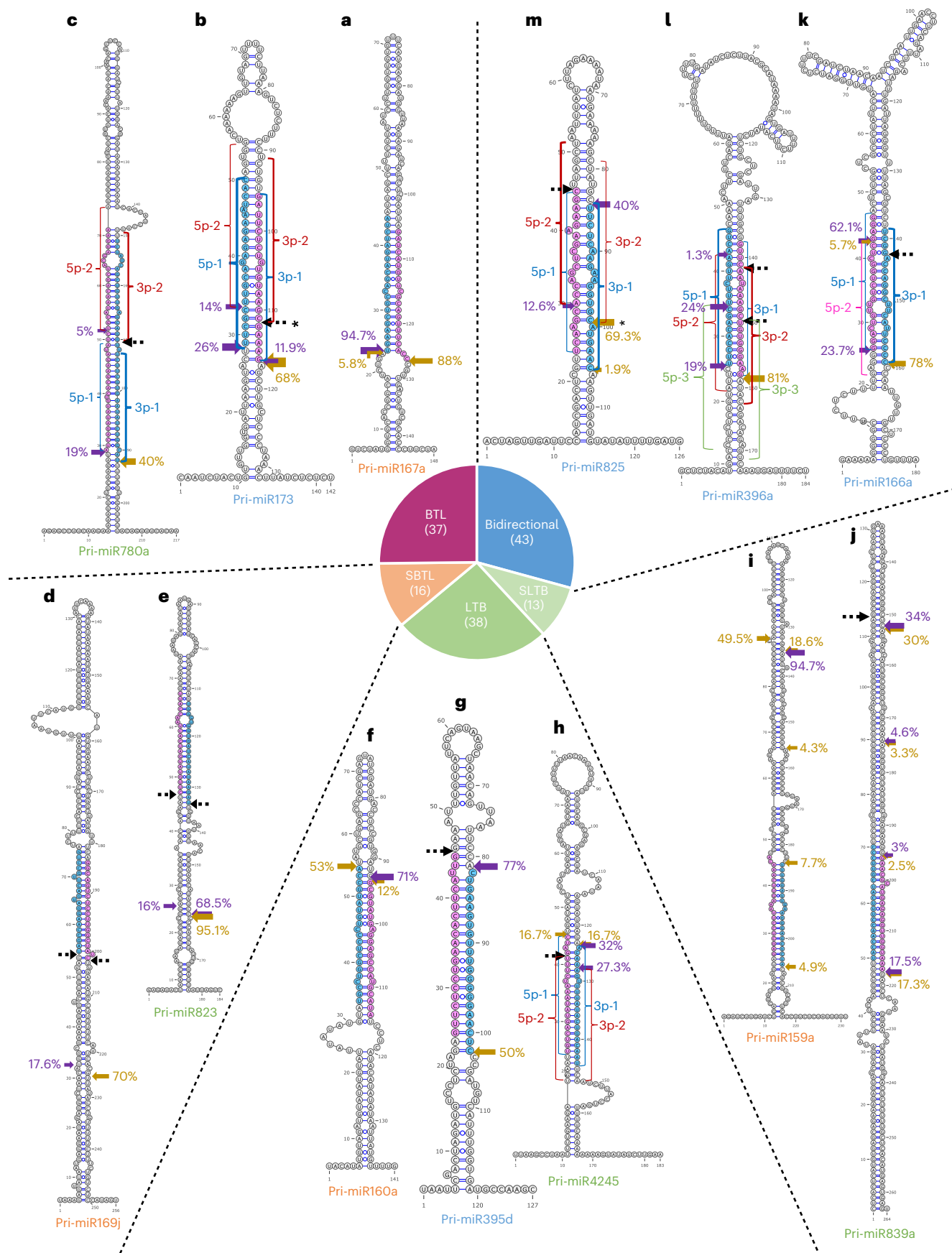
Pri-miRNAs (37) with the BTL pattern. We found that 37 of 147 pri-miRNAs adhered to the BTL processing mode (Fig. 2a–c and Supplementary Fig. 2). Among these, 21 pri-miRNAs were previously demonstrated^{12,14,15} and were re-validated unambiguously here (Fig. 2a and Supplementary Fig. 2a,b(i)–(iii)). These pri-miRNAs include pri-miR161; pri-miR165a; pri-miR166f and pri-miR166g; pri-miR167a, pri-miR167c and pri-miR167d; pri-miR168b; pri-miR169a; pri-miR170; pri-miR172b; pri-miR393b; pri-miR395b and pri-miR395f; pri-miR397a and pri-miR397b; pri-miR399a, pri-miR399d and pri-miR399f; pri-miR827; and pri-miR864.

Pri-miR173, pri-miR399b and pri-miR156f were re-annotated to the BTL pattern. Notably, pri-miR173 predominantly generates the canonical miR173/* as reported previously^{14,15} (Fig. 2b and Supplementary Fig. 2c). However, we observed an additional cutting site with a 5-nt shift at a considerable ratio (14%), leading to a new species of miRNA/* that is clearly enriched in AGO2 complexes^{29,30} (Supplementary Fig. 2d). In line with this, a target (AT4G20460) of the new miRNA could be readily recovered from our degradome-seq data (Supplementary Fig. 2e). We thus designated the new species of miRNA/* as 5p-2 in the 5' arm and 3p-2 in the 3' arm (red brackets), whereas we refer to the previously annotated products as 5p-1 in the 5' arm and 3p-1 in the 3' arm (blue brackets) (Fig. 2b). Similarly, pri-miR399b and pri-miR156f also exhibited two distinct starting sites for cleavage and generated two pairs of functional miRNA/*s (5p-1/3p-1 in blue and 5p-2/3p-2 in red; Supplementary Fig. 2f,g).

We also report the BTL processing pattern for 13 species of pri-miRNAs whose processing directions were unknown. They are pri-miR158b, pri-miR776, pri-miR777, pri-miR780a, pri-miR831, pri-miR847, pri-miR851a, pri-miR857a, pri-miR859, pri-miR867, pri-miR4240, pri-miR5013 and pri-miR5020b (Fig. 2c and Supplementary Fig. 2h–j(i),(ii)). Most of these pri-miRNAs showed clear and

Fig. 2 | Degradome-seq reveals that 37, 16, 38, 13 and 43 pri-miRNAs display the BTL, SBTL, LTB, SLTB and bidirectional processing patterns, respectively. **a**, Pri-miR167a is 1 of 21 examples whose canonical BTL processing patterns are fully validated here. **b**, Pri-miR173 is one of three examples whose processing patterns are re-annotated here. The cleavage site marked with an asterisk was also discovered in *fiery1* (ref. 14). **c**, Pri-miR780a is 1 of 13 examples whose processing patterns are newly identified. **d**, Pri-miR169j is 1 of 13 examples whose canonical SBTL processing patterns are fully validated here. **e**, Pri-miR823 is one of three newly identified examples following the SBTL processing pattern. **f**, Pri-miR160a is 1 of 13 examples whose canonical LTB processing patterns are fully validated here. **g**, Pri-miR395d is one of five examples that are revised to be processed through LTB. **h**, Pri-miR4245 is 1 of 20 examples whose processing patterns are newly identified to follow the LTB direction. **i**, Pri-miR159a is one of five examples whose canonical SLTB processing patterns are fully validated here. **j**, Pri-miR839a is one of eight newly identified examples following the

SLTB pattern. **k**, Pri-miR166a is 1 of 24 canonical pri-miRNAs with bidirectional processing patterns. **l**, Pri-miR396a is one example of five re-annotated bidirectional pri-miRNAs with new productive and abortive products. **m**, Pri-miR825 is one example of 14 re-annotated bidirectional pri-miRNAs with new productive products. It can follow BTL processing to produce a non-canonical but abundant pair of miRNA/* (5p-2/3p-2) besides the previously reported LTB pattern that produces canonical miR825/* (5p-1/3p-1). The cleavage site marked with an asterisk was also discovered in the WT¹⁴. In **a–m**, the percentages indicate the relative ratios of cleavage sites of individual pri-miRNAs by two semi-active DCL1 variants. Black dotted arrows denote the intended cutting positions that are not detectable in our system. Thicker arrows indicate higher cutting ratios. The pri-miRNA names in orange, blue and green indicate re-validated, re-annotated and newly identified ones, respectively. The five patterns (BTL, SBTL, LTB, SLTB and bidirectional) are summarized in the middle.



unique first cleavage sites at the base regions and generated uniform pre-miRNAs for the next round of processing. However, there were some exceptions. Pri-miR780a showed two initial cutting sites proximal to the base region: the predominant one, which produced a large amount of miR780/* (5p-1/3p-1, in blue), and a minor one that generated a new species of miRNA/* (5p-2/3p-2, in red). Both products were functional, because sufficient read counts could be recovered from AGO1-RISC (Fig. 2c and Supplementary Fig. 2h,i). Similarly, three other pri-miRNAs, pri-miR831, pri-miR847 and pri-miR851a, also displayed two or more types of the BTL processing pattern and new functional or abortive forms of miRNA/*s (Supplementary Fig. 2j(i)).

Pri-miRNAs (16) with the SBTL pattern. Certain pri-miRNAs have extended stems and undergo sequential cleavages before reaching the positions of miRNA/*s. Approximately 13 pri-miRNAs have been reported to exhibit the SBTL processing pattern^{12–15} and were re-validated here. These pri-miRNAs include pri-miR163; pri-miR169b, pri-miR169c, pri-miR169f, pri-miR169g, pri-miR169i, pri-miR169j, pri-miR169l, pri-miR169m and pri-miR169n; pri-miR394a; pri-miR402; and pri-miR447b (Fig. 2d and Supplementary Fig. 3a,b(i),(ii)). We also extended the list to include pri-miR823, pri-miR822 and pri-miR845a (Fig. 2e and Supplementary Fig. 3c,d).

Pri-miRNAs (38) with the LTB pattern. A total of 13 pri-miRNAs have been reported to undergo the LTB mode of processing^{12,14,15} and were fully validated in our study (Fig. 2f and Supplementary Fig. 4a,b(i),(ii)). These pri-miRNAs include pri-miR156a–pri-miR156e and pri-miR156h, pri-miR160a–pri-miR160c, pri-miR162a, pri-miR171b and pri-miR171c, and pri-miR400. We re-annotated the LTB processing pattern for pri-miR395d, pri-miR159c, pri-miR390b, pri-miR408 and pri-miR779a (Fig. 2g and Supplementary Fig. 4c,d).

We deciphered the LTB processing mode for 20 new species of pri-miRNAs (Fig. 2h and Supplementary Fig. 4e,f(i)–(iii)). They include pri-miR775, pri-miR833a, pri-miR841a, pri-miR848a, pri-miR849a, pri-miR852, pri-miR853, pri-miR1888a, pri-miR2112, pri-miR3440b, pri-miR4245, pri-miR5012, pri-miR5017, pri-miR5021, pri-miR5026, pri-miR5028, pri-miR5634, pri-miR5637, pri-miR5640 and pri-miR5654. A representative case is pri-miR4245, which produced two sets of miRNA/*s: one was the canonical 5p-1/3p-1 (in blue), while the other was the newly identified 5p-2/3p-2 (in red); they could both be recovered from AGO-RISC (Fig. 2h and Supplementary Fig. 4e). Similarly, pri-miR775, pri-miR848a, pri-miR3440b and pri-miR5017 (Supplementary Fig. 4f(i),(ii)) all showed two or three kinds of the LTB processing pattern that generated new species of miRNA/*s in addition to the previously reported miRNA/*s. Notably, the processing site for the canonical miR848/* (5p-1/3p-1, in blue) was not detectable in our study, although very low reads of the miRNA/* could be recovered from AGO1-RISC. In contrast, the new set of miRNA/* (5p-2/3p-2, in red) can be recovered from AGO1-RISC even with more reads (Supplementary Fig. 4f(i)).

Pri-miRNAs (13) with the SLTB pattern. Five pri-miRNAs, pri-miR159a and pri-miR159b and pri-miR319a–pri-miR319c, were previously reported to have the SLTB pattern and were further validated in our study (Fig. 2i and Supplementary Fig. 5a,b). We also extended this pattern to eight new pri-miRNAs including pri-miR835a, pri-miR839a, pri-miR840a, pri-miR856, pri-miR869a, pri-miR1888b, pri-miR5024 and pri-miR5656 (Fig. 2j and Supplementary Fig. 5c,d). Interestingly, pri-miR869a could produce a new set of miRNA/* (5p-2/3p-2, in red) that was even more abundant than the annotated 5p-1/3p-1 (in blue) (Supplementary Fig. 5d). Pri-miR1888b produced two sets of miRNA/*s (5p-1/3p-1, in blue; 5p-2/3p-2, in red) which could be detected in sRNA-seq data. However, neither of them is identical to the predicted miRNA/*, which should be re-annotated (Supplementary Fig. 5d). Additionally, pri-miR5656 did not produce the annotated duplex (5p-1/3p-1, in blue) but produced a new miRNA/* product, highlighted

by red brackets (5p-2/3p-2), which can be recovered from AGO1-RISC (Supplementary Fig. 5d).

Pri-miRNAs (43) with the bidirectional processing pattern. Many plant pri-miRNAs showed the bidirectional processing pattern due to the presence of terminal branched loops⁶. One example is pri-miR166a, which could undergo BTL processing to produce the canonical miR166/* (5p-1/3p-1) or an LTB mode to produce an abortive product, 5p-2 (Fig. 2k and Supplementary Fig. 6a). In fact, this scenario could be extended to an additional 23 pri-miRNAs (Supplementary Fig. 6b(i)–(iv)). Surprisingly, many founding pri-miRNAs such as pri-miR158a, pri-miR164b and pri-miR164c, pri-miR165b, and pri-miR166b displayed bidirectional processing. These pri-miRNAs do not have terminal branched loops, but they have big internal loops that serve as triggers for bidirectional processing (Supplementary Fig. 6b(ii)).

Pri-miR396a displayed a unique pattern of processing (Fig. 2l and Supplementary Fig. 6c): it underwent major BTL processing and produced the annotated miR396/* (5p-1/3p-1, in blue) that was loaded into AGO1-RISC. This pri-miRNA could also be processed in an LTB direction, yielding a new set of productive products (5p-2/3p-2, in red). The miR396/* (5p-2/3p-2) was sorted into AGO2 with considerable read numbers and had a new target (AT2G29340) for silencing (Supplementary Fig. 6d,e). Moreover, this pri-miRNA could also be cut right in the middle of the miR396/* duplex with a moderate frequency (24%) but yielded a set of abortive products (5p-3/3p-3, in light green) (Fig. 2l). This pattern with three initial processing sites was also detected for pri-miR162b, pri-miR172c, pri-miR167b and pri-miR844a (Supplementary Fig. 6f(i),(ii)).

We also found that 14 pri-miRNAs could employ a bidirectional processing pattern but produce two sets of miRNA/*s that are both functional. For instance, pri-miR825 could be cleaved in an LTB direction, producing the annotated miR825/* (blue bracket) (Fig. 2m and Supplementary Fig. 6g). This pri-miRNA could also be processed in a BTL manner, producing the new functional 5p-2/3p-2, which was sorted into AGO2 (red brackets) (Fig. 2m and Supplementary Fig. 6h). The other 13 pri-miRNAs with such a pattern included pri-miR157c, pri-miR169d, pri-miR171a, pri-miR393a, pri-miR396b, pri-miR398a, pri-miR403, pri-miR472a, pri-miR824, pri-miR868a, pri-miR3933, pri-miR5014a and pri-miR8183 (Supplementary Fig. 6i(i)–(iv)).

Taken together, we have validated the previously reported processing patterns for 52 pri-miRNAs^{12–15,31}, while re-annotating the processing modes for 43 pri-miRNAs and reporting the processing patterns of 52 new pri-miRNAs that were not revealed before (Fig. 2 and Supplementary Table 1). We sorted all 147 pri-miRNAs into five categories according to their cleavage patterns and present these new productive miRNA/*s as well as their potential targets in Supplementary Table 2.

SE and HYL1 act in defining the first cutting sites of pri-miRNAs

We next assessed the influence of HYL1 and SE on the determination of the initial processing sites on pri-miRNAs. To this end, we re-sorted the 147 bona fide pri-miRNAs into three patterns: BTL-, LTB- and sequential-processed pri-miRNAs (including SBTL and SLTB; Methods). Since the sites with the highest cutting ratios in the *DCL1*^{E1507Q} and *DCL1*^{E1696Q} lines represent the first cleavage positions in pri-miRNAs, we designated positions as ‘O’ for all pri-miRNAs processed in either the BTL or LTB mode. We then extended from the ‘O’ positions to 100 nt upstream and downstream at 10-nt intervals and counted the number of initial cleavages that fell at each interval of pri-miRNAs in Col-0, *DCL1*^{E1507Q}, *DCL1*^{E1696Q}, *hyl1-2* and *se-2* (Fig. 3a,b). Importantly, ensemble analysis of all pri-miRNAs did not reveal significant accumulation of the initial cutting sites across both BTL and LTB pri-miRNAs in *hyl1-2* and *se-2*. These results suggest that the first cut selections of most pri-miRNAs are dependent on HYL1/SE (Fig. 3a,b). Alternatively, these pri-miRNAs might not be accessible to DCL1 for cleavage in the mutants.

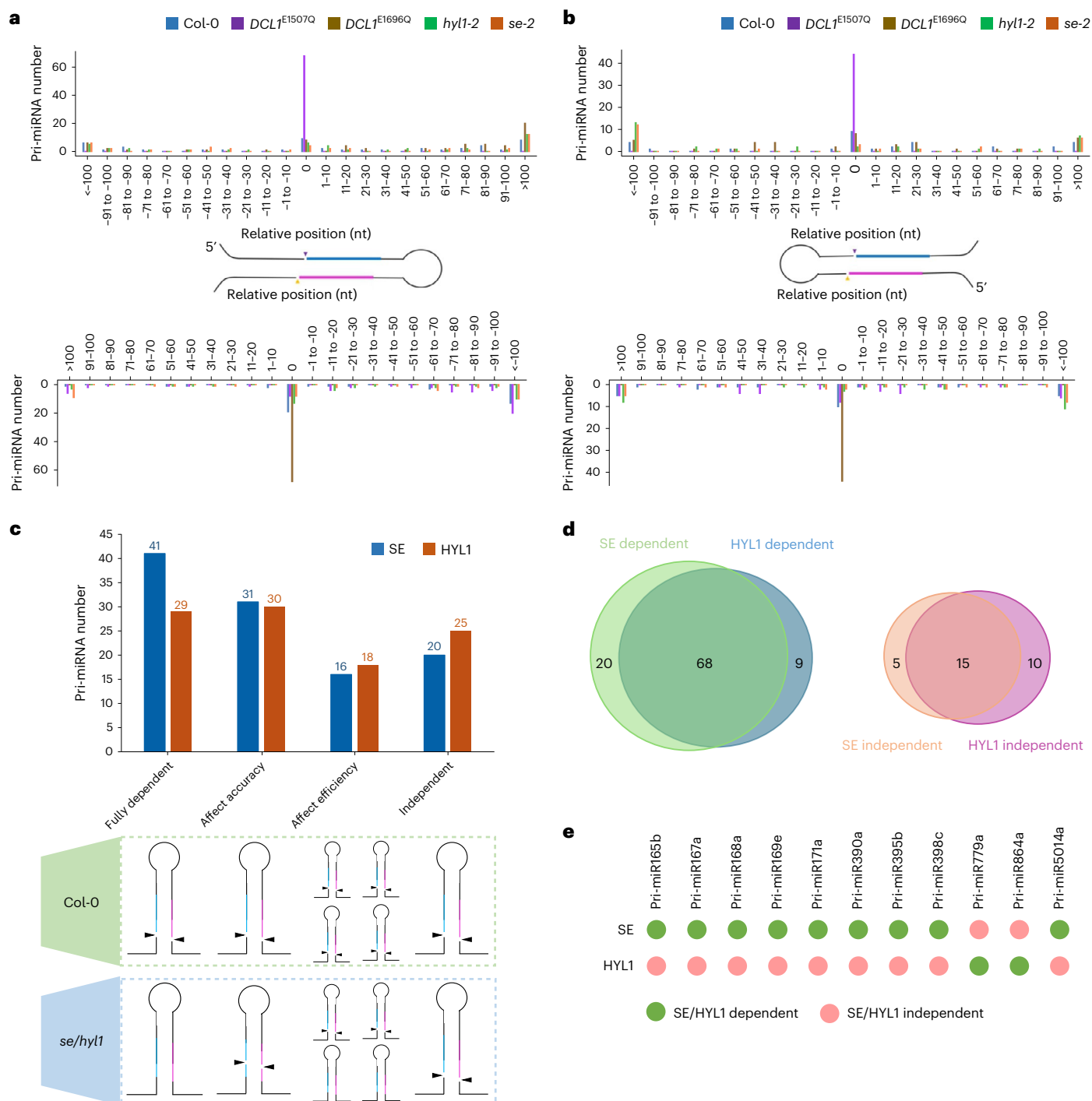


Fig. 3 | SE and HYL1 show different impacts on pri-miRNA processing.

a,b, Distributions of the first cleavage sites for 72 BTL-processed (**a**) and 46 LTB-processed (**b**) pri-miRNAs in Col-0 (blue), $DCL1^{E1507Q}$ (purple), $DCL1^{E1696Q}$ (yellow-brown), *hyl1-2* (green) and *se-2* (orange). The positions labelled '0' are defined as the first cutting sites detected in the $DCL1^{E1507Q}$ and $DCL1^{E1696Q}$ transgenic lines in the top and bottom panels and shown by the purple and yellow arrowheads, respectively, in the pri-miRNA diagrams. The blue and pink

regions in the pri-miRNA diagrams represent the miRNA/* duplex. The data are from two biological replicates for each sample. **c**, The numbers (top) and processing patterns (bottom) of pri-miRNAs that are impacted by SE and/or HYL1 proteins. Note that only the first cutting sites of BTL patterns are indicated by black arrowheads. **d**, Venn diagram showing overlapping of miRNAs that are dependent on or independent of SE and HYL1. **e**, List of 11 pri-miRNAs for which SE and HYL1 exert opposite impacts on their processing.

Certain pri-miRNAs, however, indeed exhibited the first cutting sites at position '0' in the *se* and *hyl1* mutants, implying that the absence of SE or HYL1 does not impact the selection of the first cutting sites for these pri-miRNAs (Fig. 3a,b). The numbers of such pri-miRNAs in the *se* and *hyl1* mutants are lower than those in Col-0, indicating that the selection of first cleavages for the missing pri-miRNAs can be

dependent on SE and HYL1 (Fig. 3a,b). In this scenario, we re-sorted pri-miRNAs on the basis of the impact of SE/HYL1 on the selection of the initial cleavage sites. First, certain pri-miRNAs did not exhibit any cutting events in their transcript regions in the *se* or *hyl1* mutants. These pri-miRNAs could be genuinely classified as SE/HYL1-dependent, as their initial processing fully relies on SE or HYL1. With this definition,

approximately 41 and 29 of 147 pri-miRNAs are fully dependent on SE and HYL1, respectively (Fig. 3c). By contrast, one group of pri-miRNAs (20 or 25 of 147 pri-miRNAs, respectively) exhibited identical or even higher cutting ratios at position 'O' in the *se* or *hyl1* mutants compared with those of Col-0 (Fig. 3c and Supplementary Fig. 7a,b). This pattern strongly indicates that the initial processing of these pri-miRNAs is genuinely independent of either SE or HYL1. Further analysis identified two other groups that fell between dependence on and independence of SE/HYL1. Approximately 31 or 30 pri-miRNAs could undergo the first cleavages into pri-miRNA regions in the *se* or *hyl1* mutants, but the first cutting sites could be in the lower stem, miRNA/* duplex, upper stem and even terminal loop regions for both BTL and LTB pri-miRNAs (Fig. 3a–c and Supplementary Fig. 7a,b). This pattern indicates that SE/HYL1 affects the processing accuracy of pri-miRNAs, and the absence of SE/HYL1 leads to random cutting throughout the pri-miRNAs. This pattern also suggests that SE/HYL1 might act as a molecular ruler for this type of pri-miRNAs. Finally, one group of pri-miRNAs could be also cleaved at position 'O' but with lower cutting ratios in *se/hyl1* mutants, indicative of a significant decrease in cleavage efficiency. This type of pri-miRNAs could be classified as SE/HYL1-partially-dependent, and the numbers are approximately 16 or 18 of 147 pri-miRNAs (Fig. 3a–c and Supplementary Fig. 7a,b). This group of pri-miRNAs might require relay through SE/HYL1 to DCL1 to be processed.

Next, we investigated the impact of SE/HYL1 on the initial processing of 28 pri-miRNAs following SBTL/SLTB processing patterns. By designating the first nucleotide of the miRNA/* duplex as position 'O', we observed that the processing of the pri-miRNAs relied on SE/HYL1 to different extents considering the correct selection of the initial cutting sites and high cutting ratios in the *se/hyl1* mutants compared with Col-0 (Supplementary Fig. 7c,d). Together, these findings are well in line with the previously published sRNA data^{17,32}, which demonstrate that not all miRNA levels are significantly decreased in *se* or *hyl1* mutants. Of course, the temporal/spatial expression difference between *MIR* loci and *SE/HYL1* might also contribute to the variation in the miRNAs' dependency on the proteins.

In summary, we reclassified pri-miRNAs into four categories on the basis of their initial cleavage's dependence on the two proteins: SE/HYL1-fully-dependent, SE/HYL1-dependent accuracy-affected, SE/HYL1-dependent efficiency-affected and SE/HYL1-independent (Fig. 3c). Among the 116 pri-miRNAs with enough detectable reads, 41, 31, 16 and 20 fell into these four categories regarding SE's dependence, respectively. In parallel, 29, 30, 18 and 25 of 116 pri-miRNAs displayed varied dependences on HYL1, respectively (Fig. 3c). When combining all SE- or HYL1-fully-dependent or partially dependent pri-miRNAs, we observed that 68 pri-miRNAs are reliant on both SE and HYL1, exemplified by pri-miR398b, as their cleavage sites were not detectable at all (Fig. 3d and Supplementary Fig. 7e), whereas 15 pri-miRNAs are entirely independent of SE and HYL1, exemplified by pri-miR319a, because their processing patterns were exactly the same in *se/hyl1* mutants as in the DCL1 semi-active lines (Fig. 3d and

Supplementary Fig. 7f). Moreover, 11 pri-miRNAs showed different requirements regarding SE or HYL1 for processing (Fig. 3e). Specifically, 2 of the 11 pri-miRNAs, pri-miR779a and pri-miR864a, required HYL1 but not SE. Conversely, the remaining 9 pri-miRNAs required SE but not HYL1 for their miRNA biogenesis, including pri-miR165b, pri-miR167a, pri-miR168a, pri-miR169e, pri-miR171a, pri-miR390a, pri-miR395b, pri-miR398c and pri-miR5014a (Fig. 3e and Supplementary Table 3). The processing of these 9 pri-miRNAs showed independence of HYL1, probably due to the functional redundancy of five DRB proteins in *Arabidopsis*^{33,34}. This notwithstanding, these pri-miRNAs, whether dependent on or independent of SE or HYL1, did not exhibit a preference for cleavage patterns (BTL or LTB) (Supplementary Fig. 7g and Supplementary Table 3). In summary, our degradome-seq of the *se* and *hyl1* mutants revealed distinct impacts of SE and HYL1 on the first cutting sites for pri-miRNA processing and offers detailed insights into the molecular mechanisms underlying their contributions to the processing of specific pri-miRNAs.

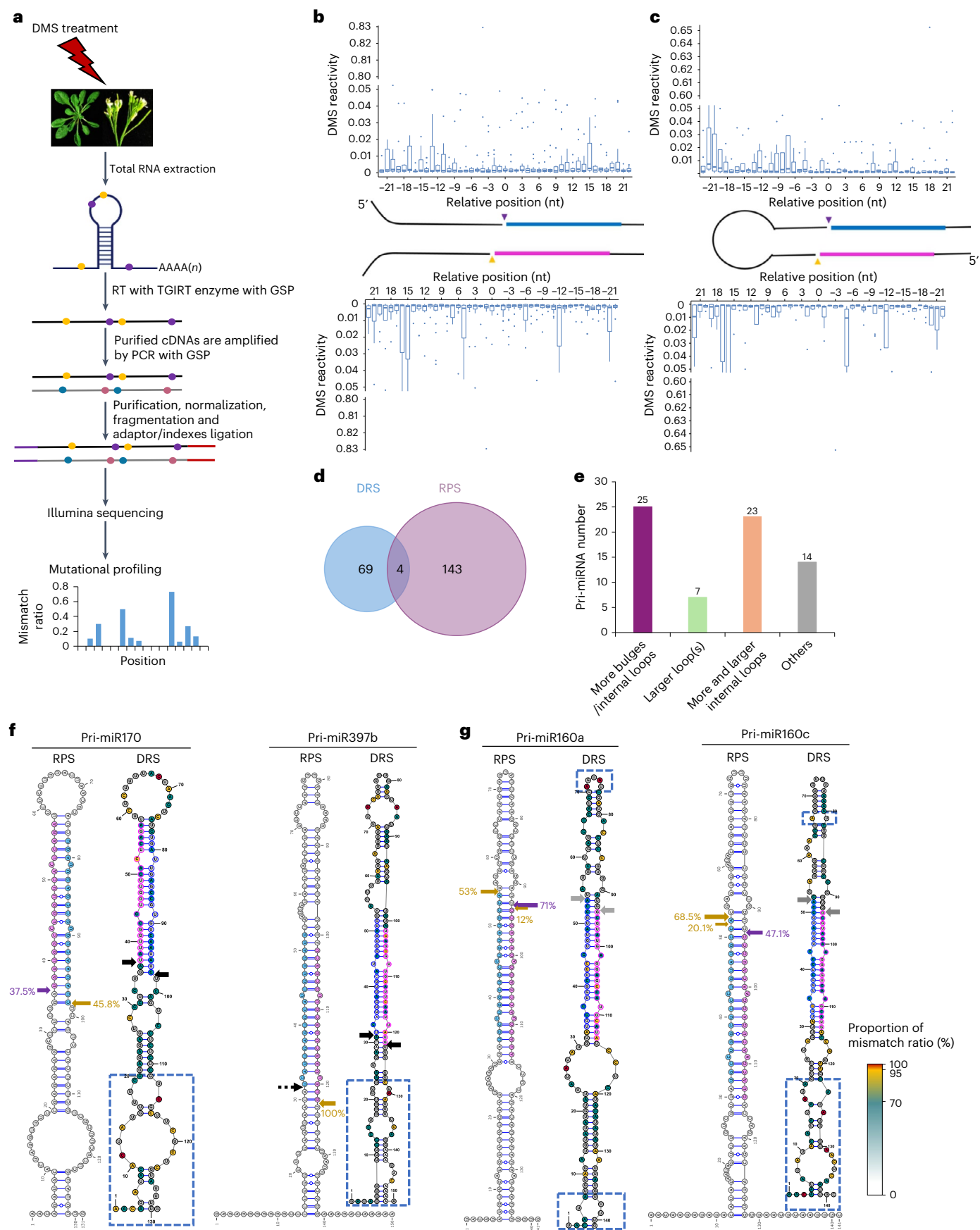
In vivo and in silico RSS of pri-miRNAs largely differ

We next investigated how the initial cutting sites and processing patterns of individual pri-miRNAs are determined in vivo. To this end, we determined the RSS of pri-miRNAs in vivo using our optimized DMS-MaPseq method²⁸. Briefly, we treated three-week-old plants and flower tissues from Col-0, *dcl1-9*, *se-1* and *hyl1-2* with dimethyl sulfate (DMS), a chemical that can methylate unpaired adenines (As) and cytosines (Cs) located at ss or loop regions of RNA molecules. The DMS-induced lesions are decoded in reverse transcription by the thermostable group II intron reverse transcriptase (TGIRT) enzyme, which can induce random mismatches for the methylated As and Cs (Fig. 4a). Bioinformatic analysis revealed that all biological replicates of Col-0, *dcl1-9*, *hyl1-2* and *se-1* largely exhibited consistently higher mismatch ratios of As and Cs than their mock-treated counterparts, indicating the reliability of the experimental strategy (Extended Data Fig. 1a).

For 147 bona fide pri-miRNAs, we were able to obtain sufficient read counts (≥ 500 reads) for 73, 77, 76 and 78 pri-miRNAs in Col-0, *dcl1-9*, *se-1* and *hyl1-2*, respectively, but not for the rest of the pri-miRNAs due to primer specificity and/or extremely low expression levels. First, 22 nt were extended both upstream and downstream from the first cleavage sites (position 'O'), which covers the lower stem regions for the BTL-type pri-miRNAs and the upper stem regions for the LTB-type ones, as well as the miRNA/* duplex regions. Interestingly, for the BTL-patterned pri-miRNAs, the distal regions of the lower stems, or the basal regions, exhibited higher DMS reactivities than those of the duplex regions (Fig. 4b). This pattern indicates the presence of more unpaired nucleotides in the distal regions of the lower stems, which is in line with the structural features of pri-miRNAs. Conversely, for the LTB-patterned pri-miRNAs, the upper stems and terminal loop regions displayed significantly higher DMS reactivities than those of the duplex regions, indicative of single-strandedness in the upper regions of pri-miRNAs (Fig. 4c). The duplex regions of LTB-patterned pri-miRNAs

Fig. 4 | DMS-MaPseq reveals bona fide RSS of pri-miRNAs in vivo. **a**, Schematic illustration of DMS-MaPseq to detect RSS of pri-miRNAs. GSP, gene-specific primers. **b,c**, The DMS reactivities on basal segments, lower stems and duplex regions of BTL-processed pri-miRNAs (**b**) and on terminal loops, upper stems and duplex regions of LTB-processed pri-miRNAs (**c**) in Col-0. Positions labelled 'O' are defined as the first cutting sites in *DCL1*^{E1507Q} and *DCL1*^{E1696Q} transgenic lines in the top and bottom panels and are indicated by purple and yellow arrowheads, respectively, in the pri-miRNA diagrams. The blue and pink regions represent the miRNA/* duplex. The data are from three biological replicates. The central lines in the boxes represent the median values, the upper and lower bounds show the first and third quartiles, the whiskers indicate data within 1.5× the interquartile range of both quartiles, and points past the ends of the whiskers represent outliers. **d**, Venn diagram showing the overlap between DMS-reactivity-based structures (DRS) and RNAfold-predicted structures (RPS). **e**, Pri-miRNAs with

different structures in vivo and in silico are further divided into four categories on the basis of detailed structure differences. **f,g**, BTL-processed pri-miR170 and pri-miR397b (**f**) and LTB-processed pri-miR160a and pri-miR160c (**g**) are four representative examples of the 69 pri-miRNAs that show different structures between in vivo DMS-reactivity-based modelling (right) and in silico modelling (left). The annotated miRNA/* regions are shaded with the outlines in blue and pink. The black dotted arrow denotes an intended cutting position that was not detectable in our system. The black and grey arrows indicate the first cleavage sites for BTL- and LTB-patterned pri-miRNAs, respectively. The structural differences are highlighted in the blue dashed boxes. The residues with the top 5%, 25% and 70% DMS activities are labelled with red, yellow and cyan backgrounds, respectively. The remaining residues with mismatch ratios below 0.01% are labelled with a white background. Guanines (G) and uracils (U) are marked with a grey background.



tended to display higher overall DMS reactivities, indicating more open-structured duplex regions than the BTL-patterned pri-miRNAs. This pattern is also found in pri-miRNAs with sequential processing patterns (Extended Data Fig. 1b).

Next, we modelled the *in vivo* RSS of pri-miRNAs in Col-0 samples on the basis of an algorithm that combines both minimum free energy and DMS reactivity as constraints. We found that ~5.5% of pri-miRNAs (4 of 73) exhibited the same structures between DMS-reactivity-based structures (DRS) and RNAfold-predicted structures (RPS) (Fig. 4d). These pri-miRNAs were pri-miR156a, pri-miR168a, pri-miR844a and pri-miR856a (Extended Data Fig. 1c). In contrast, the vast majority, ~94.5% (69 of 73) demonstrated distinct structures between DRS and RPS. The most common differences involved the presence of more and larger internal loops or bulges in DRS than in RPS (Fig. 4d,e, Extended Data Figs. 2–8 and Supplementary Table 4). For instance, BTL-type pri-miR170 and pri-miR397b harboured three internal loops (bulges) in the basal region according to DRS, instead of one big internal loop as predicted from *in silico* RNAfold (Fig. 4f). LTB-processed pri-miR160a showed a larger terminal loop and one additional internal loop at its basal region in DRS compared with RPS (Fig. 4g, left). Another LTB-cleaved example, pri-miR160c, obtained one additional internal loop (mismatch) in the top region and four internal loops (bulges) in the basal region in DRS, instead of having one big bulge in RPS (Fig. 4g, right). Collectively, these findings highlight a substantial disparity between the *in silico* and *in vivo* structures that could help us understand the complexity of pri-miRNA processing *in vivo*.

In vivo RSS better explains the first cutting sites of pri-miRNAs

Since Microprocessor tends to cut 15–17 nt away from the ssRNA–dsRNA reference junctions, and since a vast majority of plant pri-miRNAs seem to lack such structures *in silico*, we performed a systemic survey of RSS of DMS-reacted pri-miRNAs *in vivo*. Indeed, there were internal loops or bulges in both BTL- and LTB-patterned pri-miRNAs, and their distances from the first cleavage sites peaked at 15–17 nt with roughly normal distributions (Fig. 5a).

Systemic assessment of all 73 pri-miRNAs with detectable *in vivo* RSS revealed that an additional ~5% of pri-miRNAs adapted to the optimized length of the molecular ruler (15–17 nt) for BTL-patterned processing on the basis of *in vivo* RSS compared with *in silico* folding (Extended Data Fig. 9a and Fig. 5a, left). For instance, RNAfold predicted that pri-miR397a displayed 6-nt and 19-nt distances from the internal loop and ss basal segment to the initial processing site, respectively, while *in vivo* RSS showed that the pri-miRNA obtained a new bulge (10 nt away), and a new internal loop (14 nt away) in the lower stem that is closer to 15 nt and thus could more efficiently guide BTL processing (Fig. 5b, left). For another BTL-type pri-miR851a, two internal loops in the lower stem were detected from *in vivo* RSS, serving as preferable reference sites (16 nt and 9 nt) for Microprocessor versus *in silico* ones (9 nt and 3 nt) (Fig. 5b, right).

For the LTB-patterned pri-miRNAs, we found that an additional ~13% contained new structural elements that could serve as more optimal molecular rulers (15–17 nt) on the basis of *in vivo* RSS compared with RPS (Extended Data Fig. 9b and Fig. 5a, right). For example, *in vivo* RSS showed that pri-miR390b now exhibited a new small bulge in the upper stem and thus introduced a 16-nt ruler to guide LTB processing. Moreover, another 10-nt distance between a new internal loop and the initial cleavage appeared in DRS. However, both the 16-nt and 10-nt lengths are obscure in the RPS (Fig. 5c, left). Another LTB-type pri-miR4245 presented a new bulge in the upper stem that served as a more optimal 15-nt reference for a new species of miRNA/* from DRS than the 20-nt distance in RPS for one set of cuttings. In addition, this same new bulge could introduce an 11-nt length to another set of cuttings, instead of the 7-nt distance in RPS (Fig. 5c, right). Taken together, approximately 5% and 13% more BTL- and LTB-patterned pri-miRNAs presented more optimal reference sites for DCL1 in *in vivo* RSS analysis than *in silico*, facilitating the activity of Microprocessor (Fig. 5b,c, Extended Data Fig. 9a,b and Supplementary Table 5).

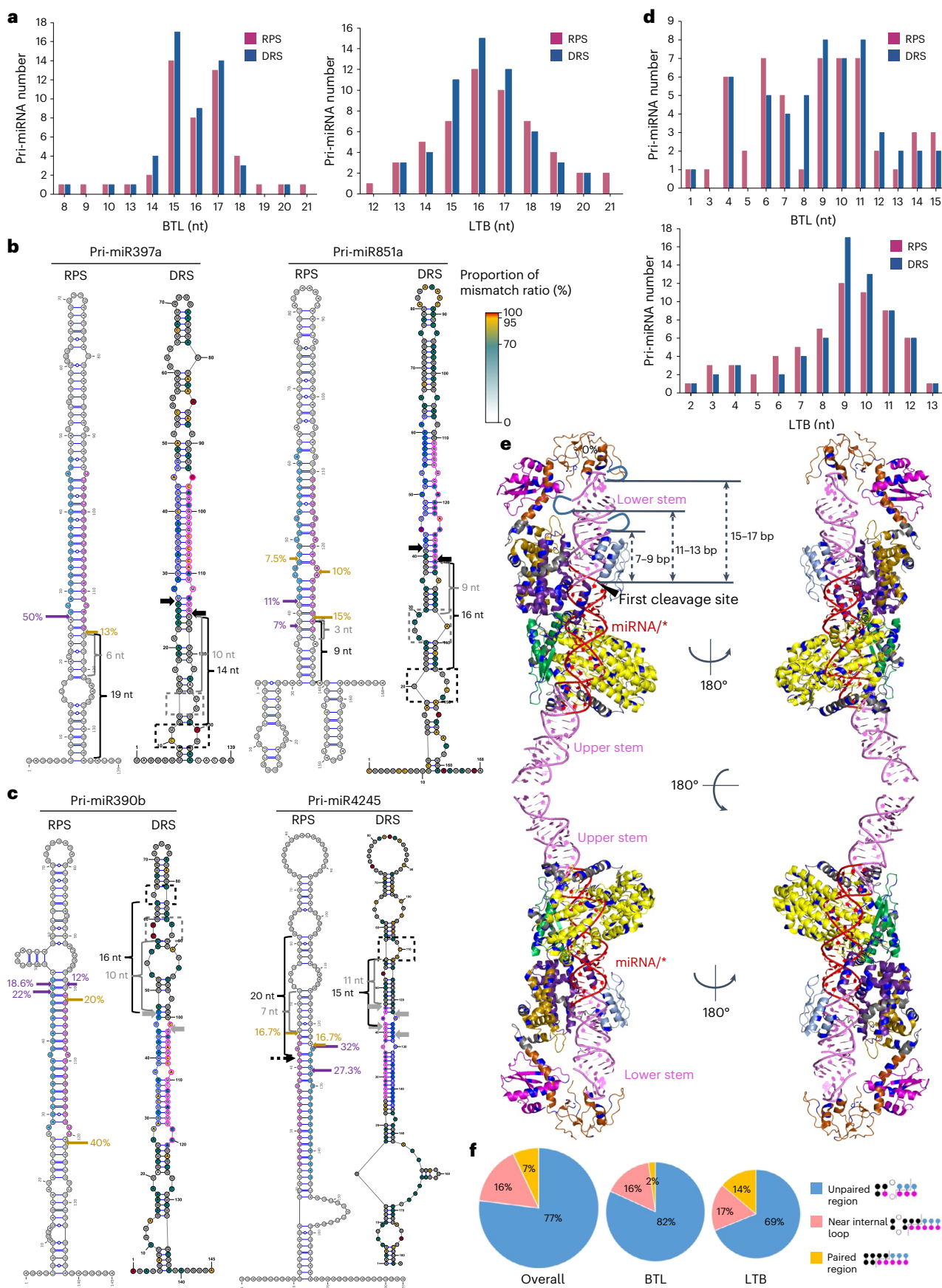
New features of internal loops/bulges and DCL1's first cuts

Intriguingly, both BTL- and LTB-patterned pri-miRNAs tended to harbour additional internal loops or bulges nearby that are approximately 9–11 nt away from the first cleavages (Fig. 5d,e). This characteristic is more pronounced in LTB-processed pri-miRNAs, as nearly the same numbers of pri-miRNAs contained the two loops/bulges that are 9–11 or 15–17 nt away from the initial processing sites with 23 species in common. Similarly, the corresponding numbers for BTL-patterned pri-miRNAs were 23 and 40, respectively, with 20 in common (Extended Data Fig. 9c). These results indicate that Microprocessor might have more flexibility to adapt to different locations of internal loops as observed *in vitro*⁶. Of course, whether additional internal loops/bulges serve as new hidden or additional molecular rulers or regulatory elements awaits future investigation.

Furthermore, we explored the structural features of the first cutting positions for all 73 pri-miRNAs with DRS in Col-0. Overall, approximately 77% of pri-miRNAs had their first cleavage sites located in an unpaired region, comprising small internal loops or bulges. Around 16% of pri-miRNAs had their first cutting sites situated near internal loops, while only 7% had their first cleavages at fully paired regions (Fig. 5f, left). Further characterization of the first cutting sites based on different processing patterns revealed clear differences between BTL- and LTB-processed pri-miRNAs. For the BTL pattern, the first cleavage sites were predominantly located in unpaired regions (82%), 16% had their first cutting sites situated near internal loops and only 2% showed their first cleavage at paired regions (Fig. 5f, middle). These results indicate that Microprocessor predominantly prefers to place the first cleavage sites at unpaired regions of BTL pri-miRNAs. By contrast, for LTB pri-miRNAs, the proportion of first cutting sites located in unpaired regions is lower (69%, right), and the proportion of first cutting sites in

Fig. 5 | DRS provides more meaningful interpretation for determination of the initial cleavages of pri-miRNAs than RPS. a, Profiling of distances from the reference dsRNA–ssRNA junction sites derived from both RNAfold and DMS-MaPseq to the first cleavage sites revealed that 15–17 nt is the predominant molecular ruler length for BTL-processed (left) and LTB-processed (right) pri-miRNAs. **b**, Pri-miR397a (left) and pri-miR851a (right) are two representative BTL-type examples that have more optimal distances for DCL1 processing between the internal reference regions and the first cleavage sites (black arrows) detected in DRS (right) than predicted from RPS (left). The black and grey brackets show 15–17-nt and 9–11-nt lengths from the reference loops, shown in black and grey dashed boxes, to the first cutting sites, respectively. **c**, Pri-miR390b (left) and pri-miR4245 (right) are two representative LTB-type examples that have more optimal distances for DCL1 processing between the internal reference regions and the first cleavage sites (grey arrows) detected in DRS (right) than predicted from RPS (left). The black dotted arrow denotes an intended cutting position

that was not detectable in our system. **d**, Both RNAfold and DMS-MaPseq showed that pri-miRNAs typically harboured extra internal loops or bulges positioned approximately 9–11 nt away from their initial cleavage sites for BTL-processed (top) and LTB-processed (bottom) pri-miRNAs. **e**, Reanalysis of the cryogenic electron microscopy density map of the DCL1–pri-miR166f complex from published data²⁵ suggests the presence of new binding pockets for additional internal loops/bulges that might be 9–11 nt away from the first cleavage sites (black arrowhead). Dark blue represents positively charged surfaces of DCL1. The blue lines indicate three different bulges in the lower stem of pri-miR166f. The colour scheme for the different domains of DCL1 is the same as in Fig. 1b. **f**, Parallel DMS-MaPseq and degradome-seq analyses show that the first cleavage sites are predominantly located at the unpaired regions for BTL-processed pri-miRNAs, whereas the pattern is less pronounced for LTB-typed pri-miRNAs. Solid and hollow circles represent paired and unpaired nucleotides, respectively. Solid blue and pink circles represent the partial miRNA/* duplex.



paired regions is higher (17% for near internal loops and 14% for paired regions) (Fig. 5f, right). Again, Microprocessor appears to have less preference in selecting the first cleavage sites for LTB-type pri-miRNAs than for BTL-type ones. These results indicate that artificial miRNA design entails more stringent selection of the first cutting site when BTL-patterned backbones are used over LTB-patterned backbones.

DCL1, SE and HYL1 act in shaping the structures of pri-miRNAs

We next assessed whether and how DCL1, SE and HYL1 impact pri-miRNAs' structures. We also computed the overall DMS reactivities for all common detectable BTL and LTB pri-miRNAs in Col-0, *dcl1-9*, *se-1* and *hyl1-2*. Interestingly, ensemble DMS reactivities for both BTL (Fig. 6a) and LTB (Fig. 6b) patterns in all three mutants were obviously elevated in comparison with Col-0. Furthermore, we calculated the Gini index for pri-miRNAs, where higher values indicate larger structure-omes. The Gini index of pri-miRNAs in *hyl1-2* and *se-1* showed a significant decrease, and that in *dcl1-9* showed a slight decrease, compared with those in Col-0 (Extended Data Fig. 10a). These results suggest that the structures of pri-miRNAs are more open in these mutants than in Col-0. Since the prevailing view is that ribonucleoprotein complexes do not seem to impact the impermeability and reactivity of DMS with bound RNA³⁵, the more open structures or heterogeneity of pri-miRNAs in the mutants might result from conformational dynamics or from lack of function of helicases that are recruited through SE or HYL1, or DCL1 itself.

Subsequently, we modelled RSS of selected pri-miRNAs in the three mutants using RNAstructure software constrained by DMS reactivity. Pri-miR395f, a representative of the BTL-processed pri-miRNAs, displayed a compact lower stem, a 2-nt internal loop in the miRNA/* duplex and two internal loops in the upper stem in Col-0. In contrast, this pri-miRNA exhibited an enlarged internal loop in the lower stem and a small but new internal loop in the proximity of miRNA/*, but no internal loop in the miRNA/* in *dcl1-9*. On the other end, two small but new internal loops occurred in *dcl1-9*, replacing one internal loop observed in Col-0. Similarly, pri-miR395f also displayed distinct structures in the *hyl1-2* and *se-1* mutants, reflected by loose-shaped upper stems among other changes in the lower stems and miRNA/* regions (Fig. 6c).

Pri-miR156a, a representative of the LTB-patterned pri-miRNAs, displayed similar folding in *dcl1-9* and Col-0 despite the seemingly increased DMS reactivity in the basal areas. This result suggests that there might be no helicase directly involved in this situation, leading to the lack of obvious structural changes in *dcl1-9*. However, this pri-miRNA clearly exhibited more unpaired regions in the basal regions and/or upper stem in *hyl1-2* and *se-1* than in Col-0 (Fig. 6d). Additional examples including SBTL-processed pri-miR447b, SLTB-processed pri-miR319a and bidirectionally processed pri-miR166a displayed varying degrees of structural changes in *dcl1-9*, *hyl1-2* and *se-1* compared with Col-0 (Extended Data Fig. 10b–d). These findings collectively suggest that DCL1, SE and HYL1 directly or indirectly modulate the RSS of pri-miRNAs.

Discussion

Understanding miRNA biogenesis has been a major challenge for plant miRNAs, as their substrates display profound heterogeneity

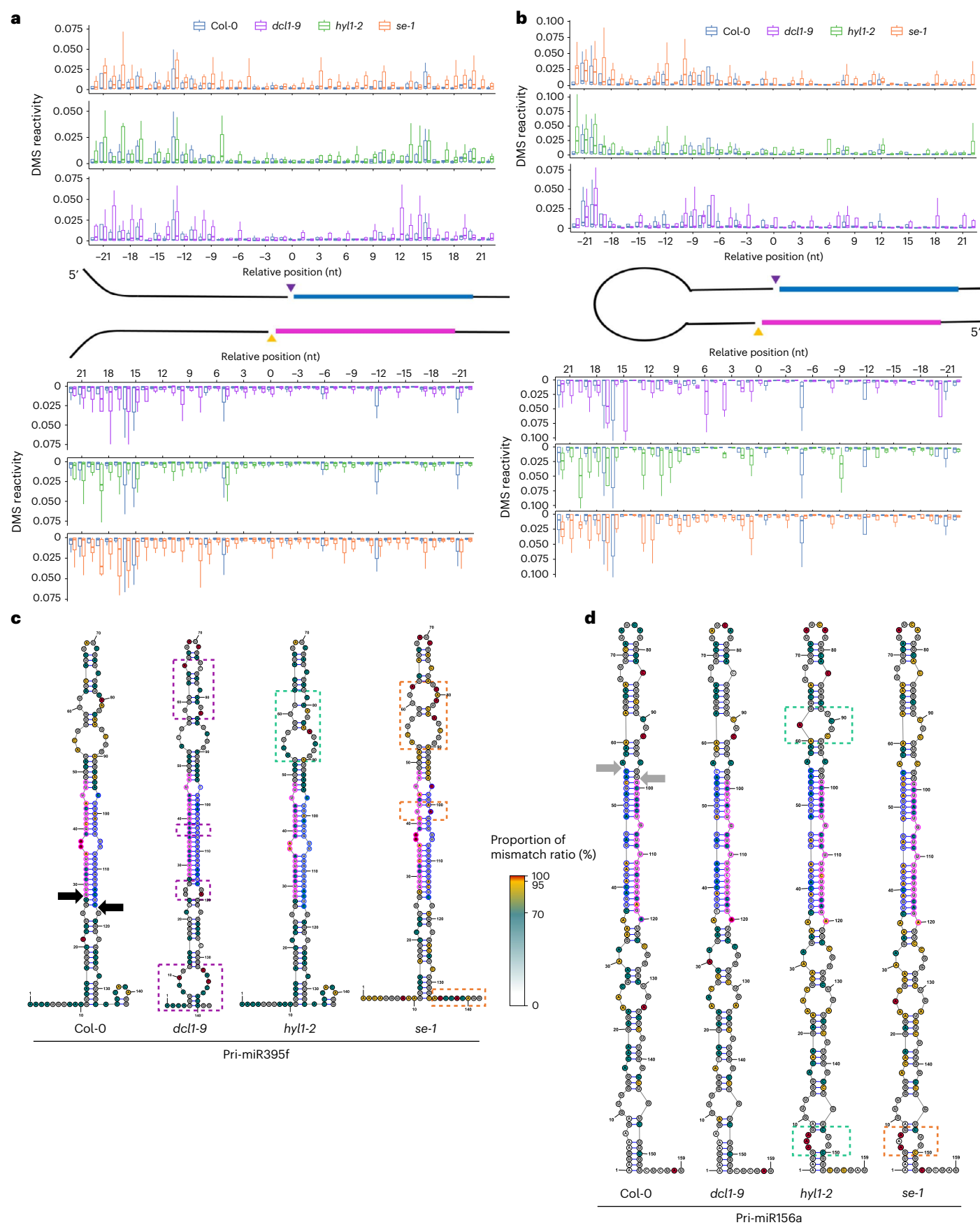
in structures and lengths^{36–38}. Here we managed to identify the precise first cutting sites for all annotated pri-miRNAs through a unique strategy of degradome-seq of DCL1 dominant-negative lines. We were also able to draw a comprehensive atlas of processing patterns of all pri-miRNAs in *Arabidopsis*. One unexpected finding is that only 147 of 326 annotated miRNAs can now be unambiguously validated as bona fide species at this stage. Importantly, we clarified that nearly 25% (81 of 326) of pri-miRNAs are not bona fide substrates of DCL1. Additionally, 98 pri-miRNAs do not have reads in degradome-seq, and their authenticity awaits future clarification (Fig. 1h).

Precise identification of the first cutting sites for all bona fide pri-miRNAs enabled us to survey their processing patterns. While we validated the processing patterns of 52 pri-miRNAs (35%) from earlier work in the field^{12,14,15,31}, we re-annotated or newly reported the processing modes for 95 pri-miRNAs (65%) (Fig. 2). Systematic investigation of 147 true pri-miRNAs now places them into BTL, SBTL, LTB, SLTB and bidirectional processing types, whose shares are 37, 16, 38, 13 and 43, respectively (Fig. 7a). Importantly, 28 pri-miRNAs can produce new species of productive miRNA/*s as they are born by Microprocessor and loaded into AGOs. We have also pinpointed the cognate mRNA targets for 10 of the 28 newly identified miRNAs (Supplementary Figs. 2 and 6). The physiological implications of these new species of miRNAs, especially their impact on canonical miRNA homeostasis and functional connections between two sets of miRNA/*s from the same pri-miRNAs and between two sets of targets, would be exciting topics for further studies.

Furthermore, DMS-MaPseq largely re-shaped our understanding of plant pri-miRNA processing from both structural and sequence perspectives. First, in vivo RSS exhibited significant differences from the in silico predicted structures. These differences are underscored by the fact that approximately 95% of 73 pri-miRNAs with detectable RSS demonstrate a prevalence of larger or more loops/bulges in vivo than the in vitro RPS (Figs. 4 and 7b). This observation is reminiscent of human pre-miRNAs that also display larger terminal loops in icSHAPE-MaP-determined RSS than in the theoretical model from miR-Base³⁹. Like their animal counterparts, plant pri-miRNAs possess more unpaired nucleotides in their RSS in vivo. Second, approximately 78% and 68% of BTL- and LTB-patterned pri-miRNAs are processed through the canonical molecular ruler (15–17 nt). These in vivo RSS features largely clarified ~5% additional BTL- and ~13% additional LTB-processed cleavage patterns that could not be previously explained on the basis of RNAfold structures. This notwithstanding, we also noticed that approximately 43% of BTL- and 61% of LTB-type pri-miRNAs harbour additional internal loops and bulges that are ~9–11 nt away from the initial cleavages (Fig. 7b and Extended Data Fig. 9a,b). These additional internal loops and bulges seem to fit into the new pockets of DCL1 found through the reanalysis of DCL1's cryogenic electron microscopy structure with pri-miR166f (Fig. 5e)²⁴. DCL1 clearly uses a binding pocket in the PAZ domain to recognize the canonical reference internal loops of pri-miRNA that are 15–17 nt away from the cleavage site. However, DCL1 also harbours an ample and adaptable space that could accommodate the internal loops/bulges that are 9–11 nt away from the cleavage site (Fig. 5e). This DCL1 structural feature underscores our observation that pri-miRNAs exhibit more and larger internal loops or bulges in DRS than in RPS (Fig. 4d–g and Extended Data Figs. 2–8). Third, we

Fig. 6 | DCL1, SE and HYL1 maintain the proper pri-miRNA secondary structures for processing. **a,b**, DMS reactivities on the basal segments, lower stem and duplex regions for BTL-processed pri-miRNAs (**a**) and on the terminal loops, upper stem and duplex regions for LTB-processed pri-miRNAs (**b**) in Col-0 (blue), *dcl1-9* (purple), *hyl1-2* (green) and *se-1* (orange). Positions labelled 'O' are defined as the first cutting sites in DCL1^{E1507Q} and DCL1^{E1696Q} transgenic lines in the top and bottom panels and are shown by the purple and yellow arrowheads, respectively, in the pri-miRNA diagrams. The blue and pink regions in the pri-miRNA diagrams represent the miRNA/* duplex. The data are from 67

commonly detected pri-miRNAs from 3 biological replicates for Col-0, *dcl1-9* and *hyl1-2*, but 2 biological replicates for *se-1*. The central lines in the boxes represent the median values, the upper and lower bounds show the first and third quartiles, and the whiskers indicate data within 1.5× the interquartile range of both quartiles. **c,d**, Pri-miR395f (**c**) and pri-miR156a (**d**) are two examples of BTL-processed and LTB-processed pri-miRNAs, respectively, that show different RSS in *dcl1-9*, *hyl1-2* and *se-1* compared with Col-0. The dashed boxes indicate structural differences in the mutants.



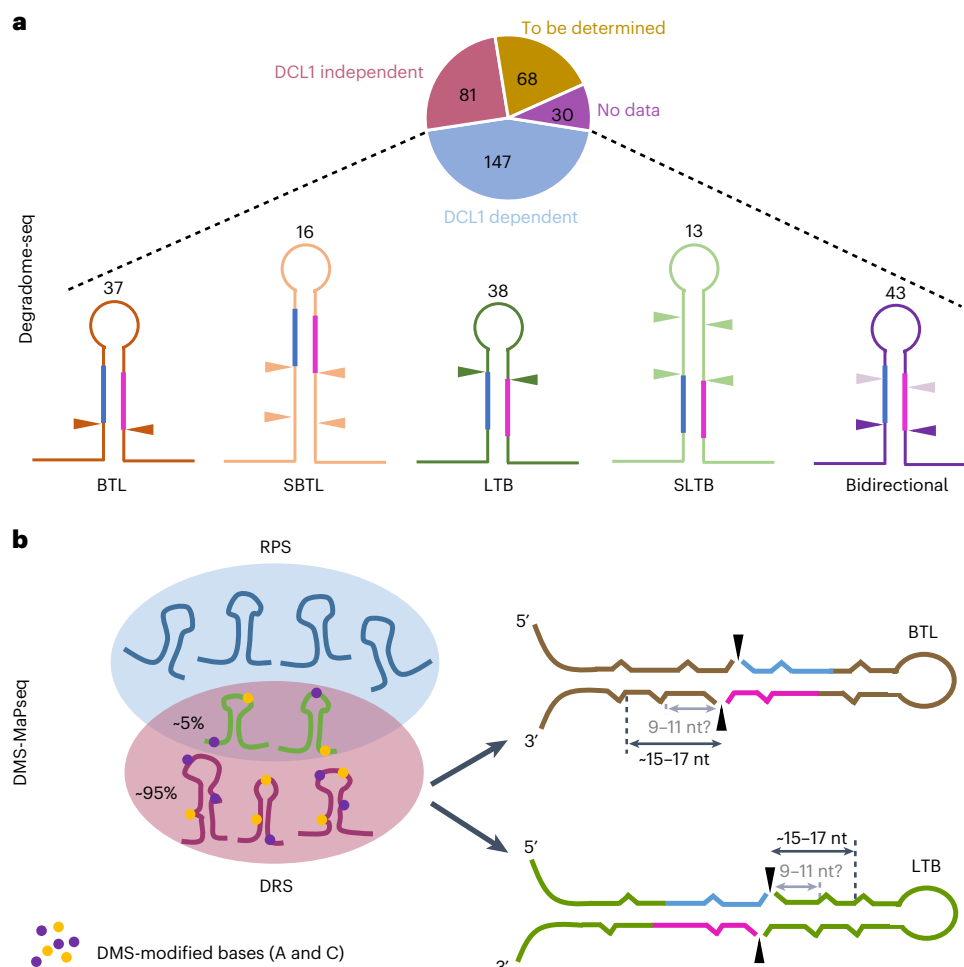


Fig. 7 | Atlas of miRNA biogenesis in *Arabidopsis* drawn from degradome-seq and DMS-MaPseq. **a**, Degradome-seq identified 147 bona fide pri-miRNAs from 326 previously annotated pri-miRNAs, and reclassified them into 5 processing patterns, namely, BTL, SBTL, LTB, SLTB and bidirectional processing. **b**, Ninety-five percent of in vivo RSS for pri-miRNAs, derived from our DMS-MaPseq (DRS),

were different from RNAfold-predicted structures (RPS). The DRS better explains why DCL1 selects the first cutting sites (black arrows) that are 15–17 nt (black lines) away from the internal loops or bulges. DRS also detects additional internal loops or bulges that are 9–11 nt (grey lines) away from the first cleavage sites by DCL1. See ‘Discussion’ for details.

observed that up to 77% of pri-miRNAs had their first cleavage sites at unpaired regions (Figs. 5f and 7b). This ratio is significantly higher than the earlier reported number (approximately 40%)⁴⁰, although the preference of DCL1 to process unpaired regions in the stem of pri-miRNAs has been appreciated before^{6,41}. These discoveries have provided new guidelines for designing artificial miRNAs: first, it will be more effective if the cleavage sites targeted by DCL1 are located on the internal unpaired regions; and second, the backbones of LTB-type pri-miRNAs seem to tolerate more mismatches or unpaired sequences in their duplex regions than the BTL-type ones. Taking pri-miR159a as an example^{42,43}, we propose to modify an amiR-backbone that resembles its DRS instead of its RPS (Extended Data Fig. 10e).

Parallel degradome-seq and DMS-MaPseq enabled us to revisit the role of SE and HYL1 in pri-miRNA processing in a physiological context. Whereas SE and HYL1 concertedly impact the initial cleavage sites, they each have independent functions. We discovered that SE and HYL1 have completely opposite effects on 11 pri-miRNAs, which had not been reported before (Fig. 3e). HYL1-independent pri-miRNAs are more abundant than SE-independent ones, and this discrepancy may be offset by the presence of HYL1 homologues in plants. Furthermore, we did not observe the preference of SE or HYL1 for the processing of specifically patterned pri-miRNAs (such as BTL or LTB). It has been well established that SE and HYL1 impact miRNA biogenesis through various

genetic and biochemical pathways^{21–23,25,26}. We have recently reported that SE could recruit SWI2/SNF2 ATPase CHR2 to remodel pri-miRNA RSS to inhibit miRNA production¹⁷. Here we can strengthen the idea that RSS can serve as a critical regulatory layer to control pri-miRNA processing. This concept can be highlighted by the observation that pri-miRNAs in *dcl1-9*, *se-1* and *hyl1-2* displayed higher DMS reactivity than those in Col-0, indicative of flexible RSS in vivo (Fig. 6 and Extended Data Fig. 10). This result further suggests that RSS changes might also contribute to the inaccurate or abnormal processing of pri-miRNAs observed in *se* and *hyl1* mutants. It is generally assumed at this stage that protein–RNA interaction negligibly impacts DMS activities on bound substrates^{35,44,45}. In this scenario, the alteration of RSS by DCL1 can be attributed to its inherent helicase function, while the impact of SE and HYL1 on RSS is probably indirect and occurs through the activities of their interacting helicases exemplified by CHR2. Otherwise, the current dogma that protein binding does not affect the chemical probing of RNA structures in vivo needs to be revisited.

In summary, parallel degradome-seq and DMS-MaPseq have provided a comprehensive atlas of pri-miRNA processing steps and the structural determinants for such processing. The dissection of connections between RSS features and first cleavage sites provides new guidelines for designing artificial miRNAs for more precise and efficient silencing of targeted transcripts. As a side outcome, these

findings unambiguously clarify the identities of the bona fide miRNAs. The ideas and approaches in this study can thus be readily adopted to different organisms.

Methods

Plant materials and growth conditions

Arabidopsis thaliana ecotype Columbia (Col-0), *dcl1-9* (CS3828), *se-1* (CS3257), *se-2* (SAIL_44_G12) and *hyl1-2* (SALK_064863) were used for this study as described previously⁴³. The *pBA002a-P_{DCL1}-FM-gDCL1^{E1507Q}* and *pBA002a-P_{DCL1}-FM-gDCL1^{E1696Q}* constructs were transformed into the *dcl1-9^{-/-}* background and screened by genotyping. The primers used for PCR identification are listed in Supplementary Table 6. All plants were grown on soil (Jorjy Gardener/LP5) or standard Murashige and Skoog (PhytoTech Labs, catalogue number M524) medium plates in a 16-hour light/8-hour dark cycle at 22 ± 1 °C.

Vector construction

The *pBA002a-P_{DCL1}-FM-gDCL1^{E1507Q}* and *pBA002a-P_{DCL1}-FM-gDCL1^{E1696Q}* used for this study were constructed as follows: *pBA-Flag-Myc-DCL1^{E1507Q}* and *pBA-Flag-Myc-DCL1^{E1696Q}* (described previously⁶) were digested by BamHI (New England BioLabs, catalogue number R3136S)/PmeI (New England BioLabs, catalogue number R0560S), and the resultant fragments were ligated into the BamHI/PmeI-treated *pBA002a-P_{DCL1}-FM-gDCL1* (ref. 17) to yield the *pBA002a-P_{DCL1}-FM-gDCL1^{E1507Q}* and *pBA002a-P_{DCL1}-FM-gDCL1^{E1696Q}* constructs; these were confirmed by sequencing using seq-F: AAGGATGAGGCAACAATGG and seq-R: ACCTTCCAAGCAGCAGTTGT. They were then transformed into the *dcl1-9^{-/-}* background via the floral dip transformation method⁴⁶.

Western blot assays

Western blot analysis was performed with ten-day-old plants that were grown on Murashige and Skoog plates or three-week-old plants that were grown on soil. The extraction and experimental procedures were performed as described previously⁴⁷. The blots were detected with primary antibodies against Myc (Sigma-Aldrich, catalogue number C 3956, 1:10,000 dilution) and actin (Sigma-Aldrich, catalogue number A0480, 1:10,000 dilution). The secondary antibodies were anti-rabbit immunoglobulin G (Cytiva, catalogue number NA934v, 1:5,000 dilution) and anti-mouse immunoglobulin G (Cytiva, catalogue number NA931v, 1:5,000 dilution).

sRNA blot assays

Total RNAs from ten-day-old seedlings were separated in 15% denaturing Urea-PAGE gels, followed by semi-dry transferring onto Hybond-N+ hybridization membranes (GE Healthcare, RPN303B). The membranes were UV crosslinked and hybridized with miRNA probes labelled with [γ -³²P] ATP (PerkinElmer), with U6 serving as a loading control. The membranes were covered with a phosphor imaging plate (GE Healthcare), and signals were detected with a Typhoon FLA7000 (GE Healthcare) as described previously⁴⁷. The probes used for sRNA blot are listed in Supplementary Table 6.

Library construction for degradome-seq

The RNA was extracted from three-week-old or flower tissues of Col-0, *dcl1-9^{-/-}*; *DCL1^{E1507Q}*, *dcl1-9^{-/-}*; *DCL1^{E1696Q}*, *hyl1-2* and *se-2* using TRIzol reagent (Sigma-Aldrich, catalogue number T9424) according to the manufacturer's instructions. We isolated poly(A)-RNA from total RNA using oligo(dT) dynabeads; then, an amount of 5 µg of DNase-treated poly(A)-RNA was ligated to the 5' RNA adaptor (rGrUrUrCrArGrArGrUrUrCrUrArCrArGrUrCrCrGrArCrGrArUrC) using T4 RNA ligase 1 (New England BioLabs, catalogue number M0204S) according to the manufacturer's instructions. The ligation products were purified with RNAClean XP beads (Beckman Coulter, reference number A63987). Then, chimeric RNA was reverse transcribed by Superscript III Reverse Transcriptase (Invitrogen, CA) using the 3' adaptor with

oligo(dT) (CCTTGGCACCCGAGAATTCCATTTTTTTTTTTTTTTTTTTVN) and amplified by PCR with adaptors including index primers (Supplementary Table 6) for Illumina sequencing.

Degradome-seq analysis

The raw data (PE150) were initially filtered for quality control requiring a quality score >25 at the 3' end and trimmed to remove the adaptor sequences by Cutadapt⁴⁸. Only Read1 reads from the PE150 data have cleavage information; these were retained for the following analysis. Clean reads over 50 nt were retained and aligned using hisat2 (ref. 49) to the *Arabidopsis* Araport11 genome reference⁵⁰. The software feature-Counts⁵¹ was used to count the read number located in pri-miRNA with a 200-bp extension in both the upstream and downstream directions. Degradome-seq recovered roughly 13–21 million clean reads for each replicate, among which more than 74% could be uniquely mapped to the reference genome. A Pearson correlation between replicates exceeded 0.92, indicating high-quality and reproducible datasets. Data from two biological replicates were merged for the following analyses due to high repeatability.

To calculate the cutting ratio for each position of all pri-miRNAs, the read count difference for each nucleotide was first determined by subtracting the read count of the upstream position from that of the current position. The cutting ratio was calculated by dividing the read count difference from the total reads mapped to the current pri-miRNA. The highest cutting ratio position was defined as the first cleavage site for each pri-miRNA, and the processing directions were sorted into five different types on the basis of the location of the first cutting site.

Prediction of the in vitro RSS of pri-miRNA via RNAfold

The secondary structures were predicted from the RNAfold web server (<http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi>) on the basis of minimum free energy and partition function. The annotated miRNA/* duplex sequences of pri-miRNAs collected from the miRBase website (<https://www.mirbase.org/browse/results/?organism=ath>) were coloured with blue and pink backgrounds, respectively, on structure models using VARNAv3-93 (<https://varna.lisn.upsaclay.fr>).

Determination of productive and abortive pri-miRNA products

We downloaded and repurposed previously published databases, which include regular sRNA-seq and AGOs-IP sRNA-seq data from both our laboratory and others^{29,30,52}. To ascertain whether these newly discovered products are functional (productive) or non-functional (abortive), we established the following criterion: if there are sufficient read counts (≥20) supporting these new products, they are categorized as productive processing. If the read counts are not evident (<20), these products are considered abortive cleavage.

Target prediction for new species of productive miRNAs

The command-line version of psRNATarget⁵³ (<https://github.com/jtremblay/mirnatarget>) was used to predict targets for 28 newly discovered miRNAs against the Araport11 cDNA reference (<https://www.arabidopsis.org>). Predicted target regions with no more than one mismatch in the seed region were preserved for subsequent validation. Cutting sites from Col-0 degradome-seq were employed to ensure the prediction accuracy. Only predicted target regions overlapping with cutting sites (with read coverage ≥20 and cutting ratio ≥0.1) were considered as authentic targets of the new miRNAs.

Assessing the impact of SE and HYL1 on pri-miRNAs' first cutting sites

The 147 bona fide pri-miRNAs were re-sorted into three patterns: 72 BTL-processed pri-miRNAs, 46 LTB-processed pri-miRNAs and 28 sequential-processed pri-miRNAs (including SBTL and SLTB). Pri-miR169d was excluded from the classification due to the presence

of two cutting ratio sites with equal proportions (50%), which could introduce ambiguity in our computational analysis. All 42 bidirectionally processed pri-miRNAs were re-assigned to either (S)BTL (34 of the cases) or LTB (8 of the cases) depending on the directions the canonical miRNAs are produced by. The positions with the highest cutting ratios in the *DCL1*^{E1507Q} and *DCL1*^{E1696Q} lines were designated as 'O' for all pri-miRNAs processed in either the BTL or LTB mode. The first cutting sites of RNase b, reflected in the *DCL1*^{E1507Q} line, are centred at position 'O' in the top panel of Fig. 3a, while the initial cutting sites of RNase a, reflected in the *DCL1*^{E1696Q} line, are illustrated in the bottom panel of Fig. 3a. These results are well aligned with our in vitro results that the RNase b and a domains process the 5' and 3' arms of BTL-patterned pri-miRNAs, respectively⁶. By contrast, an inverted pattern was shown for LTB-patterned pri-miRNAs, in which the RNase b and a domains preferably cleave the 3' and 5' arms of pri-miRNAs, reflected by the locations of the highest cutting ratios in the *DCL1*^{E1507Q} and *DCL1*^{E1696Q} lines, respectively (Fig. 3b).

Pri-miRNA-specific DMS-MaPseq library preparation

For three-week-old plants, about 5 g of Col-0, *dcl1-9*, *hyl1-2* and *se-1* plants grown on soil were collected within one hour and completely covered in 40 ml of 1× DMS reaction buffer (40 mM HEPES pH 7.5, 100 mM KCl and 0.5 mM MgCl₂) in a clean 50 ml Corning tube. DMS (Sigma, catalogue number D186309) was added to a final concentration of 1% as described previously²⁸. Mock treatment was performed by adding the same volume of deionized water. The samples were treated in the DMS reaction buffer or mock solution at room temperature under a vacuum for 15 min. To stop the reaction, a final concentration of 20% β-mercaptoethanol (Sigma-Aldrich, catalogue number M6250) was added, and the mixture was incubated for 5 min under vacuum. After being washed three times with 50 ml of cold deionized water, the samples were immediately frozen with liquid N₂ and ground into powder.

For flower tissues, the same plants were grown on soil until flowering, and then samples were collected and treated with or without DMS using the same conditions as above except the incubation time under vacuum. Different time courses for DMS treatment including 14 min, 28 min, 42 min and 56 min were also tested; eventually, 56 min was chosen for the assay after comparison, because incubation times less than 56 min are not sufficient to distinguish the DMS-treated samples from the mock-treated samples, and total RNA appeared to start decaying by 56 min.

Total RNA was isolated using TRIzol reagent according to the manufacturer's instructions and then treated with TURBO DNase (Invitrogen by Thermo Fisher, catalogue number AM2239). For each sample, 2 μg of DNase-treated RNA was mixed with gene-specific reverse transcription primers (2 pmol of each primer, up to ten gene-specific primers in one reaction). Then, 1 μl of 10 mM dNTP and DEPC-H₂O was added to a total volume of 13 μl and incubated at 65 °C for 5 min, then immediately put on ice for 2 min. Then, 4 μl of 5× First-Strand buffer (250 mM Tris-HCl pH 8.3, 375 mM KCl and 15 mM MgCl₂), 1 μl of 0.1 M DTT (prepared fresh), 1 μl of RNase inhibitor (Thermo Fisher, catalogue number AM2696) and 1 μl of TGIRT-III (Ingex, catalogue number TGIRT50) were added. The mixture was incubated at 42 °C for 30 min, and then reverse transcription proceeded at 60 °C for 1.5 h. The reaction was stopped by adding 2.3 μl of 1 M NaOH and heating the mixture at 98 °C for 15 min, then immediately putting it on ice. After neutralization by 2.5 M HCl, H₂O was added to a total volume of 30 μl for each sample, followed by purification with 39 μl of AMPure XP beads (Beckman Coulter, reference number A63881), which is intended for performing DNA cleanup and efficient removal of unincorporated dNTPs, primers, salts and other contaminants.

The cleaned cDNA was precipitated and resuspended in deionized water. Pri-miRNAs were then amplified using KOD hot start DNA polymerase (Millipore Sigma, catalogue number 71086) with gene-specific primers (Supplementary Table 5). PCR bands were gel purified and

normalized according to band intensity before library construction. PCR products were mixed equally and fragmented into 50–200 bp using a sonication machine (Diagenode SA Plcoruptor) following the manufacturer's protocol. After purification using AMPure XP beads, the fragments were subjected to end repair, adenylation and adaptor ligation using Illumina adaptors (Supplementary Table 6), mainly following the published protocol. The fragments were barcoded through adaptor ligation. The ligation products were purified again by two steps of cleanup with AMPure XP beads. Next, the purified barcoded libraries were enriched by 11 cycles of PCR using KOD hot start DNA polymerase. Finally, the PCR products were cleaned using AMPure XP beads and sent for sequencing by NovaSeq PE150.

Pri-miRNA-specific DMS-MaPseq analysis

The raw fastq files were initially filtered for quality control requiring a quality score >25 in the 3' end and trimmed to remove the adaptor sequences by Cutadapt⁴⁸. To further guarantee high sequence quality, Fastq_quality_filter from the Fastx-toolkit (http://hannonlab.cshl.edu/fastx_toolkit) was used to filter reads with low quality with the parameters $q = 25$ and $p = 80$, meaning that 80% of the nucleotides had a base call accuracy of more than 99.7%. High-quality reads over 50 nt were retained and mapped to the *Arabidopsis* Araport11 genome reference⁵⁰ by tophat2 (ref. 54) with the parameters $N = 15$; read-gap-length, 10; read-edit-dist, 15; max-insertion-length, 5; max-deletion-length, 5; $g, 3'$ (ref. 35). Only uniquely mapped reads were kept for mismatch calling. A homemade Python script called CountMismatch2Bed.py (https://github.com/changhaoli/TAMU_02RSS) was used to call mismatches. DMS reactivity was calculated as the ratio between the mismatch count and the total read count for each nucleotide. The average mismatch ratios including A/C/G/U were calculated and plotted in a bar plot using the R package ggpubr (<https://cran.r-project.org/web/packages/ggpubr/index.html>).

DMS-reactivity-guided modelling of bona fide pri-miRNAs' RSS

On the basis of the DMS signal, the secondary structures of pri-miRNAs were modelled by RNAstructure (version 6.3)⁵⁵. The following threshold was used to separate adenine and cytosine bases into paired and unpaired nucleotides to produce the best-fitting model for our experimental data. For each pri-miRNA, in Col-0 samples, we set the highest 5% of DMS reactivities of As or Cs as red, which fit to the unpaired region very well; 25% mismatch ratios were coloured with yellow, mainly located in unpaired regions; and the rest with signals higher than 0.0001 were cyan. For mutants, the same pri-miRNA follows the exact same threshold constraint to model its secondary structure. On the basis of miRBase annotation, all mature miRNA sequences are outlined in blue, while miRNA/* sequences are outlined in pink.

Selection of the reference loops/bulges for molecular rulers

The canonical molecular ruler is approximately 15–17 nt; thus, the ssRNA–dsRNA junction regions that are more than 14 nt away, but close to 15–17 nt, from the first cutting sites are considered as the reference sites (related to Fig. 5a). For the potential non-canonical 9–11-nt ruler, the junction regions that are more than 8 nt but less than 14 nt from the first cleavage sites were selected in this study (related to Fig. 5d).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The degradome-seq and DMS-MaPseq data have been deposited in the NCBI Sequence Read Archive under the BioProject database with accession code [PRJNA1092576](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1092576). The *Arabidopsis* genome reference was obtained from TAIR (<https://www.arabidopsis.org>) and the NCBI Nucleotide database ([CP002684–CP002688](https://www.ncbi.nlm.nih.gov/nucleotide/)). Information on the 326

previously annotated pri-miRNAs was from the miRBase (<https://www.mirbase.org/browse/results/?organism=ath>). sRNA-seq and AGOs-IP sRNA-seq data were obtained from the NCBI website with Gene Expression Omnibus accession codes [GSE78090](#), [GSE66599](#) and [GSM707678–GSM707691](#). All other data supporting the findings of the study are present in the main text and/or the Supplementary Information.

Code availability

The code (CountMismatch2Bed.py) used for mismatch calling of DMS-MaPseq generated in this study is accessible via GitHub at https://github.com/changhaoli/TAMU_02RSS.

References

- Bartel, D. P. MicroRNAs: target recognition and regulatory functions. *Cell* **136**, 215–233 (2009).
- Song, X. W., Li, Y., Cao, X. F. & Qi, Y. J. MicroRNAs and their regulatory roles in plant–environment interactions. *Annu. Rev. Plant Biol.* **70**, 489–525 (2019).
- Li, S., Castillo-González, C., Yu, B. & Zhang, X. The functions of plant small RNA s in development and in stress responses. *Plant J.* **90**, 654–670 (2017).
- Ma, Z. & Zhang, X. Actions of plant Argonautes: predictable or unpredictable? *Curr. Opin. Plant Biol.* **45**, 59–67 (2018).
- Gebert, L. F. R. & MacRae, I. J. Regulation of microRNA function in animals. *Nat. Rev. Mol. Cell Biol.* **20**, 21–37 (2019).
- Zhu, H. et al. Bidirectional processing of pri-miRNAs with branched terminal loops by *Arabidopsis* Dicer-like1. *Nat. Struct. Mol. Biol.* **20**, 1106–1115 (2013).
- Nguyen, T. A. et al. Functional anatomy of the human Microprocessor. *Cell* **161**, 1374–1387 (2015).
- Partin, A. C. et al. Heme enables proper positioning of Drosha and DGCR8 on primary microRNAs. *Nat. Commun.* **8**, 1737 (2017).
- Jin, W., Wang, J., Liu, C. P., Wang, H. W. & Xu, R. M. Structural basis for pri-miRNA recognition by Drosha. *Mol. Cell* **78**, 423–433 (2020).
- Partin, A. C. et al. Cryo-EM structures of human Drosha and DGCR8 in complex with primary microRNA. *Mol. Cell* **78**, 411–422 (2020).
- Park, J. E. et al. Dicer recognizes the 5' end of RNA for efficient and accurate processing. *Nature* **475**, 201–205 (2011).
- Bologna, N. G. et al. Multiple RNA recognition patterns during microRNA biogenesis in plants. *Genome Res.* **23**, 1675–1689 (2013).
- Chorostecki, U. et al. Evolutionary footprints reveal insights into plant microRNA biogenesis. *Plant Cell* **29**, 1248–1261 (2017).
- Moro, B. et al. Efficiency and precision of microRNA biogenesis modes in plants. *Nucleic Acids Res.* **46**, 10709–10723 (2018).
- Gonzalo, L. et al. R-loops at microRNA encoding loci promote co-transcriptional processing of pri-miRNAs in plants. *Nat. Plants* **8**, 402–418 (2022).
- Mencia, R., Gonzalo, L., Tossolini, I. & Manavella, P. A. Keeping up with the miRNAs: current paradigms of the biogenesis pathway. *J. Exp. Bot.* **74**, 2213–2227 (2023).
- Wang, Z. et al. SWI2/SNF2 ATPase CHR2 remodels pri-miRNAs via Serrate to impede miRNA production. *Nature* **557**, 516–521 (2018).
- Kozomara, A., Birgaoanu, M. & Griffiths-Jones, S. miRBase: from microRNA sequences to function. *Nucleic Acids Res.* **47**, D155–D162 (2019).
- Meyers, B. C. et al. Criteria for annotation of plant microRNAs. *Plant Cell* **20**, 3186–3190 (2008).
- Axtell, M. J. & Meyers, B. C. Revisiting criteria for plant microRNA annotation in the era of big data. *Plant Cell* **30**, 272–284 (2018).
- Lobb, D., Rallapalli, G., Schmidt, D. D., Martin, C. & Clarke, J. SERRATE: a new player on the plant microRNA scene. *EMBO Rep.* **7**, 1052–1058 (2006).
- Han, M. H., Goud, S., Song, L. & Fedoroff, N. The *Arabidopsis* double-stranded RNA-binding protein HYL1 plays a role in microRNA-mediated gene regulation. *Proc. Natl Acad. Sci. USA* **101**, 1093–1098 (2004).
- Vazquez, F., Gascioli, V., Crété, P. & Vaucheret, H. J. C. B. The nuclear dsRNA binding protein HYL1 is required for microRNA accumulation and plant development, but not posttranscriptional transgene silencing. *Curr. Biol.* **14**, 346–351 (2004).
- Wei, X. et al. Structural basis of microRNA processing by Dicer-like 1. *Nat. Plants* **7**, 1389–1396 (2021).
- Dong, Z., Han, M. H. & Fedoroff, N. The RNA-binding proteins HYL1 and SE promote accurate in vitro processing of pri-miRNA by DCL1. *Proc. Natl Acad. Sci. USA* **105**, 9970–9975 (2008).
- 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* **12**, 206–212 (2006).
- Xie, D. et al. Phase separation of SERRATE drives dicing body assembly and promotes miRNA processing in *Arabidopsis*. *Nat. Cell Biol.* **23**, 32–39 (2021).
- Wang, Z., Wang, M., Wang, T., Zhang, Y. & Zhang, X. Genome-wide probing RNA structure with the modified DMS-MaPseq in *Arabidopsis*. *Methods* **155**, 30–40 (2019).
- Zhang, Z., Liu, X., Guo, X., Wang, X. J. & Zhang, X. *Arabidopsis* AGO3 predominantly recruits 24-nt small RNAs to regulate epigenetic silencing. *Nat. Plants* **2**, 16049 (2016).
- Barciszewska-Pacak, M. et al. *Arabidopsis* microRNA expression regulation in a wide range of abiotic stress responses. *Front. Plant Sci.* **6**, 410 (2015).
- Bologna, N. G., Mateos, J. L., Bresso, E. G. & Palatnik, J. F. A loop-to-base processing mechanism underlies the biogenesis of plant microRNAs miR319 and miR159. *EMBO J.* **28**, 3646–3656 (2009).
- Ma, Z. et al. *Arabidopsis* Serrate coordinates histone methyltransferases ATXR5/6 and RNA processing factor RDR6 to regulate transposon expression. *Dev. Cell* **45**, 769–784 (2018).
- Curtin, S. J. et al. The roles of plant dsRNA-binding proteins in RNAi-like pathways. *FEBS Lett.* **582**, 2753–2760 (2008).
- Reis, R. S., Hart-Smith, G., Eamens, A. L., Wilkins, M. R. & Waterhouse, P. M. Gene regulation by translational inhibition is determined by Dicer partnering proteins. *Nat. Plants* **1**, 14027 (2015).
- Zubrad, M. et al. DMS-MaPseq for genome-wide or targeted RNA structure probing in vivo. *Nat. Methods* **14**, 75–82 (2017).
- Merchan, F., Boualem, A., Crespi, M. & Frugier, F. Plant polycistronic precursors containing non-homologous microRNAs target transcripts encoding functionally related proteins. *Genome Biol.* **10**, R136 (2009).
- Cuperus, J. T., Fahlgrun, N. & Carrington, J. C. Evolution and functional diversification of MIRNA genes. *Plant Cell* **23**, 431–442 (2011).
- Stepien, A. et al. Posttranscriptional coordination of splicing and miRNA biogenesis in plants. *WIREs RNA* **8**, e1403 (2017).
- Luo, Q. J. et al. RNA structure probing reveals the structural basis of Dicer binding and cleavage. *Nat. Commun.* **12**, 3397 (2021).
- Rojas, A. M. L. et al. Identification of key sequence features required for microRNA biogenesis in plants. *Nat. Commun.* **11**, 5320 (2020).
- Hirata, R. et al. Unpaired nucleotides on the stem of microRNA precursor are important for precise cleavage by Dicer-like 1 in *Arabidopsis*. *Genes Cells* **27**, 280–292 (2022).
- Li, Y. et al. Degradation of SERRATE via ubiquitin-independent 20S proteasome to survey RNA metabolism. *Nat. Plants* **6**, 970–982 (2020).
- Wang, L. et al. PRP4KA phosphorylates SERRATE for degradation via 20 S proteasome to fine-tune miRNA production in *Arabidopsis*. *Sci. Adv.* **8**, eabm8435 (2022).

44. Shan, L. et al. Nucleolar URB1 ensures 3' ETS rRNA removal to prevent exosome surveillance. *Nature* **615**, 526–534 (2023).
45. Xiang, Y. et al. Pervasive downstream RNA hairpins dynamically dictate start-codon selection. *Nature* **621**, 423–430 (2023).
46. Zhang, X., Henriques, R., Lin, S. S., Niu, Q. W. & Chua, N. H. *Agrobacterium*-mediated transformation of *Arabidopsis thaliana* using the floral dip method. *Nat. Protoc.* **1**, 641–646 (2006).
47. Zhu, H. et al. *Arabidopsis* Argonaute10 specifically sequesters miR166/165 to regulate shoot apical meristem development. *Cell* **145**, 242–256 (2011).
48. Martin, M. J. E. J. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* **17**, 10–12 (2011).
49. Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* **37**, 907–915 (2019).
50. Cheng, C. Y. et al. Araport11: a complete reannotation of the *Arabidopsis thaliana* reference genome. *Plant J.* **89**, 789–804 (2017).
51. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
52. Wang, H. et al. Deep sequencing of small RNAs specifically associated with *Arabidopsis* AGO1 and AGO4 uncovers new AGO functions. *Plant J.* **67**, 292–304 (2011).
53. Dai, X., Zhuang, Z. & Zhao, P. X. psRNATarget: a plant small RNA target analysis server (2017 release). *Nucleic Acids Res.* **46**, W49–W54 (2018).
54. Kim, D. et al. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36 (2013).
55. Xu, Z. Z. & Mathews, D. H. Experiment-assisted secondary structure prediction with RNAstructure. *Methods Mol. Biol.* **1490**, 163–176 (2016).

Acknowledgements

We thank all the members of the Zhang laboratory for their help and careful proofreading of this paper. We thank the TAMU High Performance Research Computing group for supercomputing support. This work was supported by grants from NIH (no. R35GM151976), NSF MCB (no. 2139857) and the Welch Foundation (no. A-2177-20230405) to X.Z. X.Y. was partially supported by a China Scholar Council fellowship.

Author contributions

X.Z. conceived the project. K.L. initially started the project and generated genetic materials. K.L. and T.Z. contributed equally to performing degradome-seq. J.Z. and X.L. contributed to the library construction. X.Y. performed DMS-MaPseq. C.L., X.Y., Q.X. and K.L. analysed the degradome-seq data. Z.W. and A.Y. helped pinpoint partial cleavage sites for the degradome-seq data. C.L. and X.Y. analysed the DMS-MaPseq data. S.C., X.P. and J.J.C. provided guidance and intellectual input. X.Y. and C.L. wrote the initial draft of the paper. X.Z. thoroughly edited the paper, and all authors contributed to the proofreading of the paper.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at <https://doi.org/10.1038/s41477-024-01725-9>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41477-024-01725-9>.

Correspondence and requests for materials should be addressed to Xiuren Zhang.

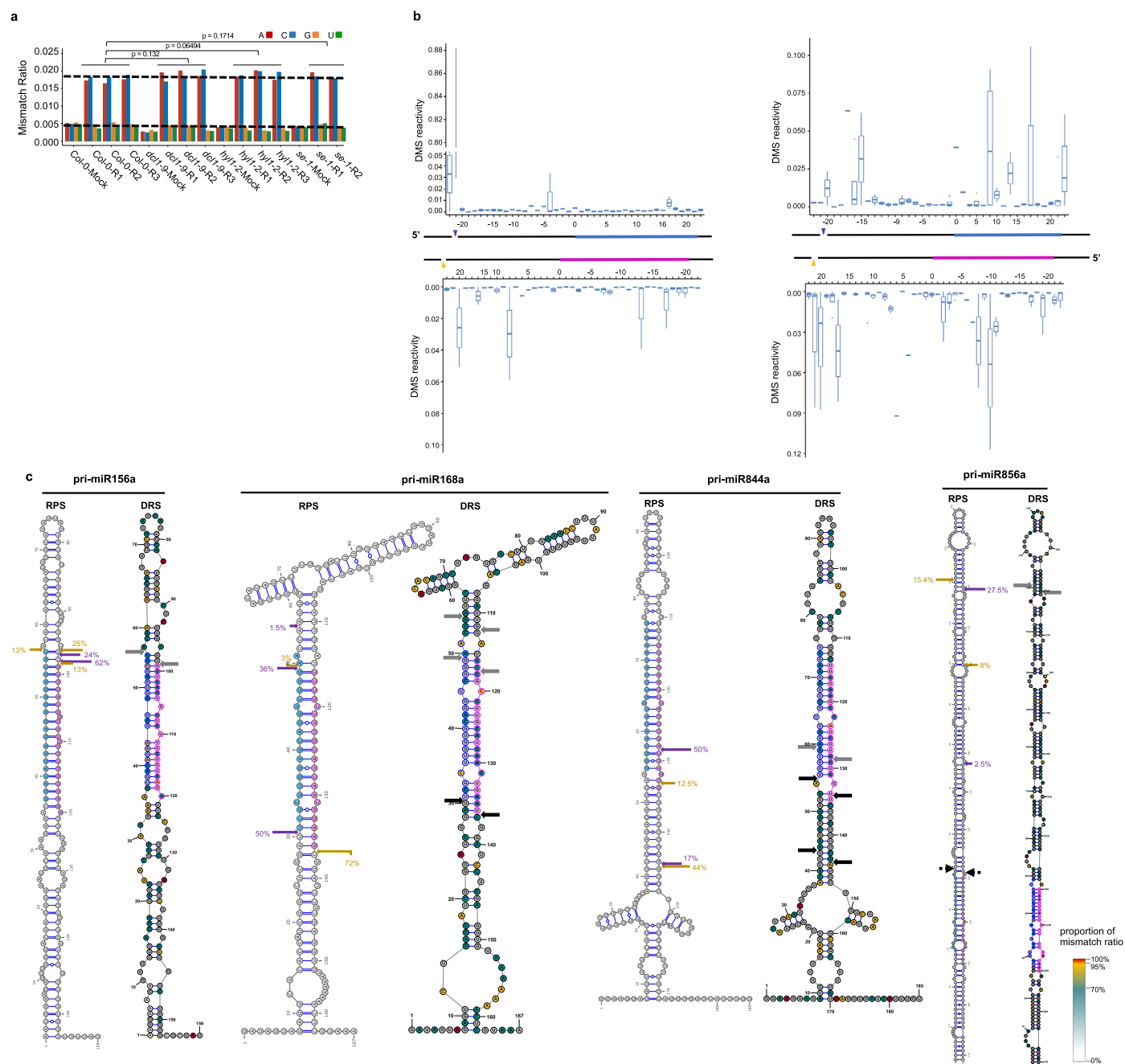
Peer review information *Nature Plants* thanks Hiro-oki Iwakawa, Peter Waterhouse and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

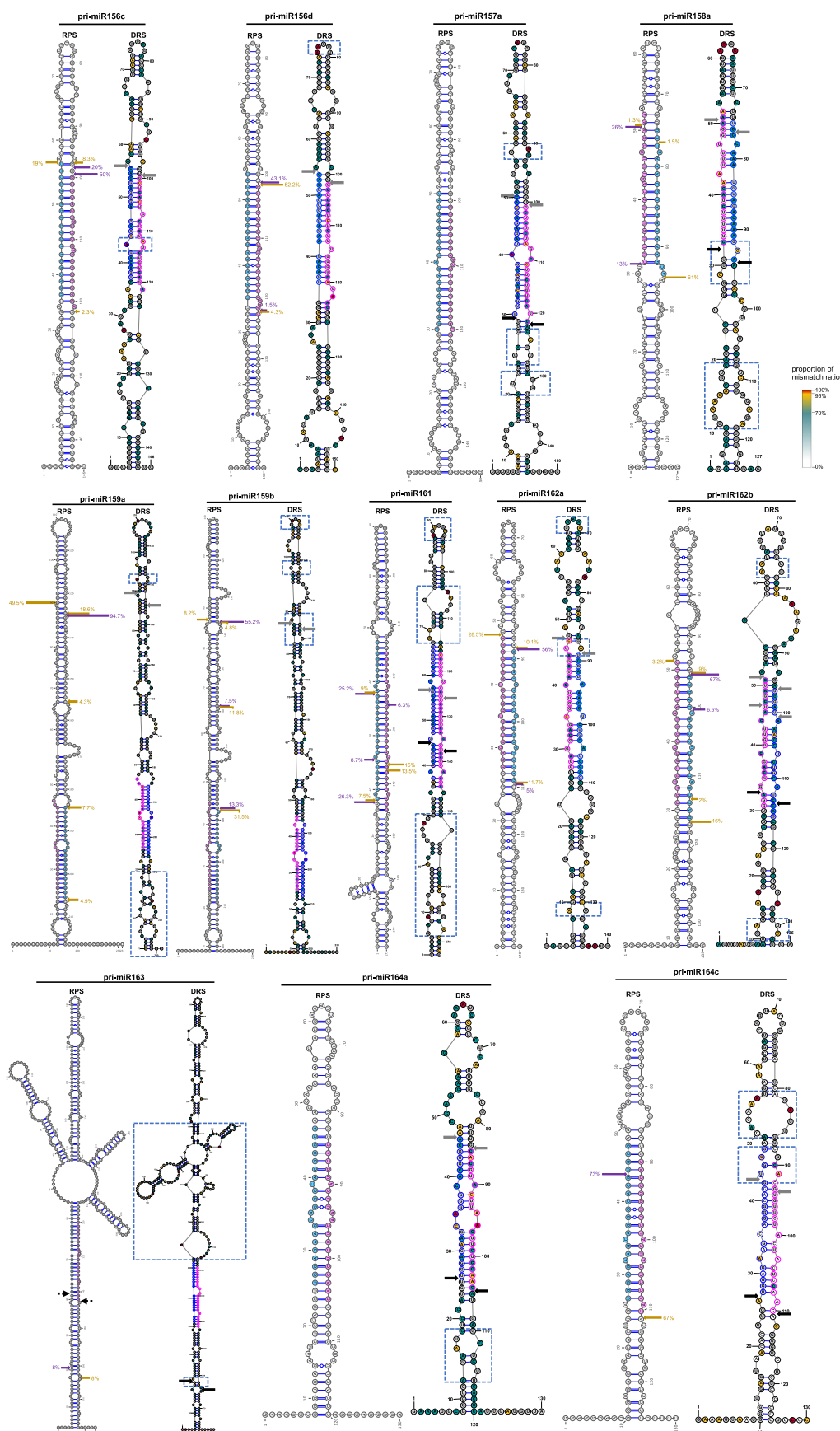
Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature Limited 2024



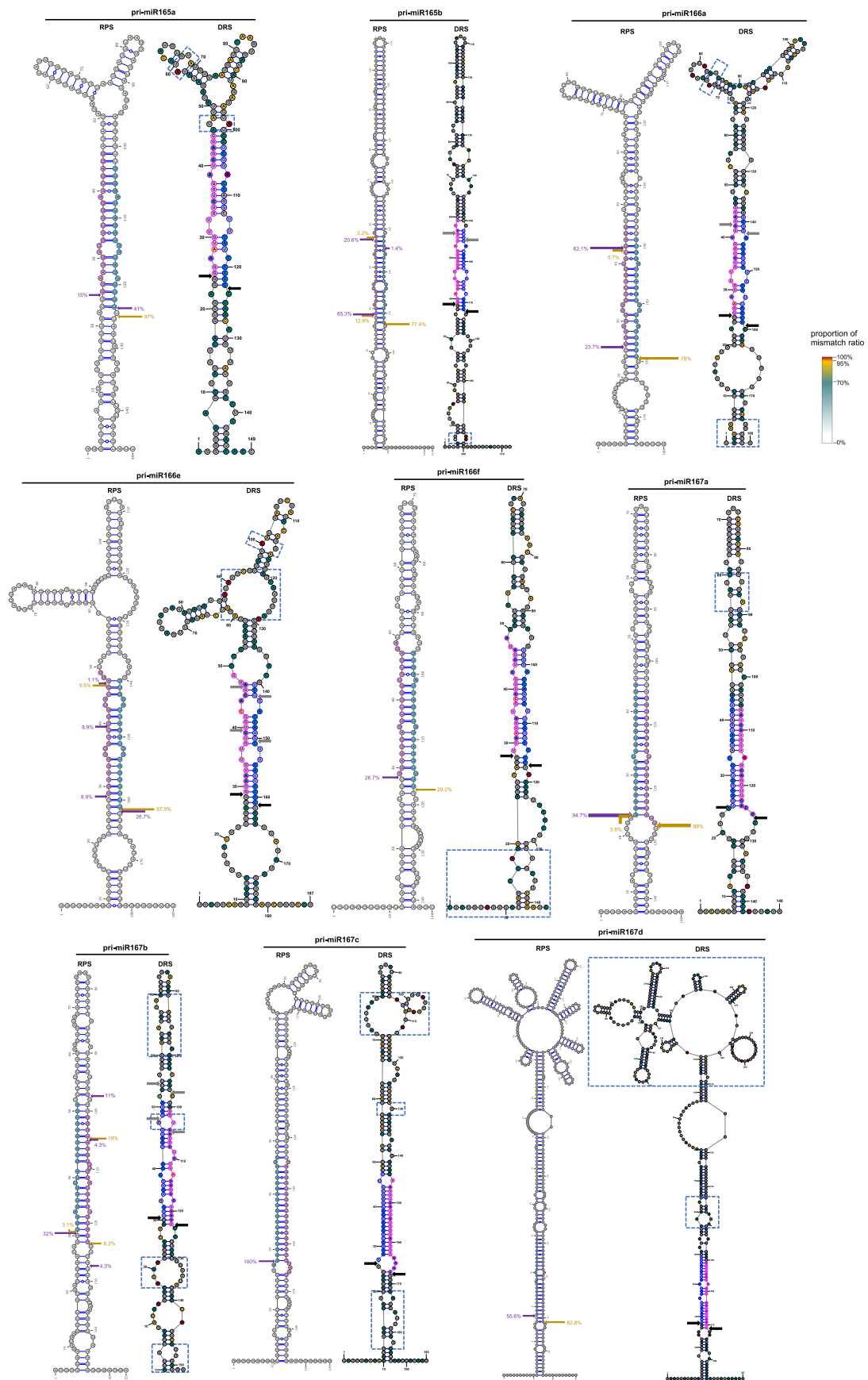
Extended Data Fig. 1 | Quality control of DMS-MaPseq library and overall patterns of DMS-MaPseq signals cross pri-miRNA backbones. (a) Average mismatch ratios of A/C/G/U caused by DMS reactivities in Col-0, *dcl1-9*, *hyl1-2* and *se-1*. The data are from 67 commonly detected pri-miRNAs from three biological replicates for Col-0, *dcl1-9* and *hyl1-2*, but two biological replicates for *se-1*. P (*dcl1-9* vs Col-0) = 0.132, P (*hyl1-2* vs Col-0) = 0.06494, P (*se-1* vs Col-0) = 0.1714. P value by Wilcoxon test. **(b)** Boxplots show the DMS reactivities for 16 SBTL (left panel) and 12 SLTB (right panel) pri-miRNAs around base/top and duplex regions in Col-0, from three biological replicates. In both top and bottom panels, position '0' is defined as the first nucleotides of duplex region, the purple

and yellow arrowheads labeled in the pri-miRNA cartoon represent the first cleavage sites. The blue and pink regions represent miRNA/* duplex. Centres of the boxes represent the median values. Upper bound and lower bound show the first and the third quartiles respectively. Whiskers indicate data within 1.5× the interquartile range of both quartiles. Data points at the ends of whiskers represent outliers. **(c)** Pri-miR156a, pri-miR168a, pri-miR844a and pri-miR856a show identical structures in DRS (right) compared to RPS (left). Black and gray arrows indicated first cutting sites for BTL and LTB directions, respectively. RPS: RNAfold Predicted Structures. DRS: DMS Reactivity based Structures.

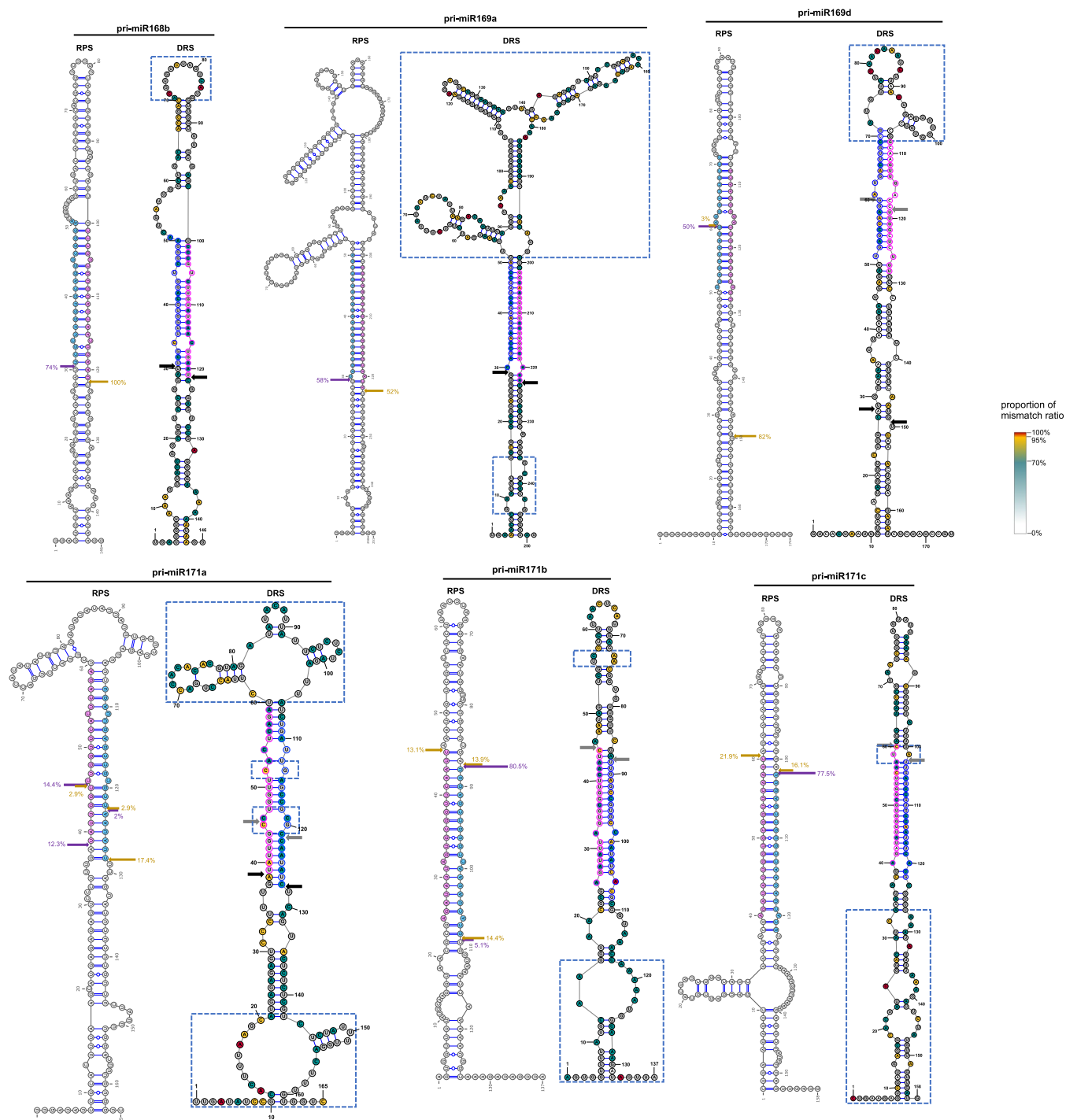


Extended Data Fig. 2 | Pri-miR156c, d, pri-miR157a, pri-miR158a, pri-miR159a, b, pri-miR161, pri-miR162a, b, pri-miR163, and pri-miR164a, c show structural differences in DRS (right) compared to RPS (left). Black and

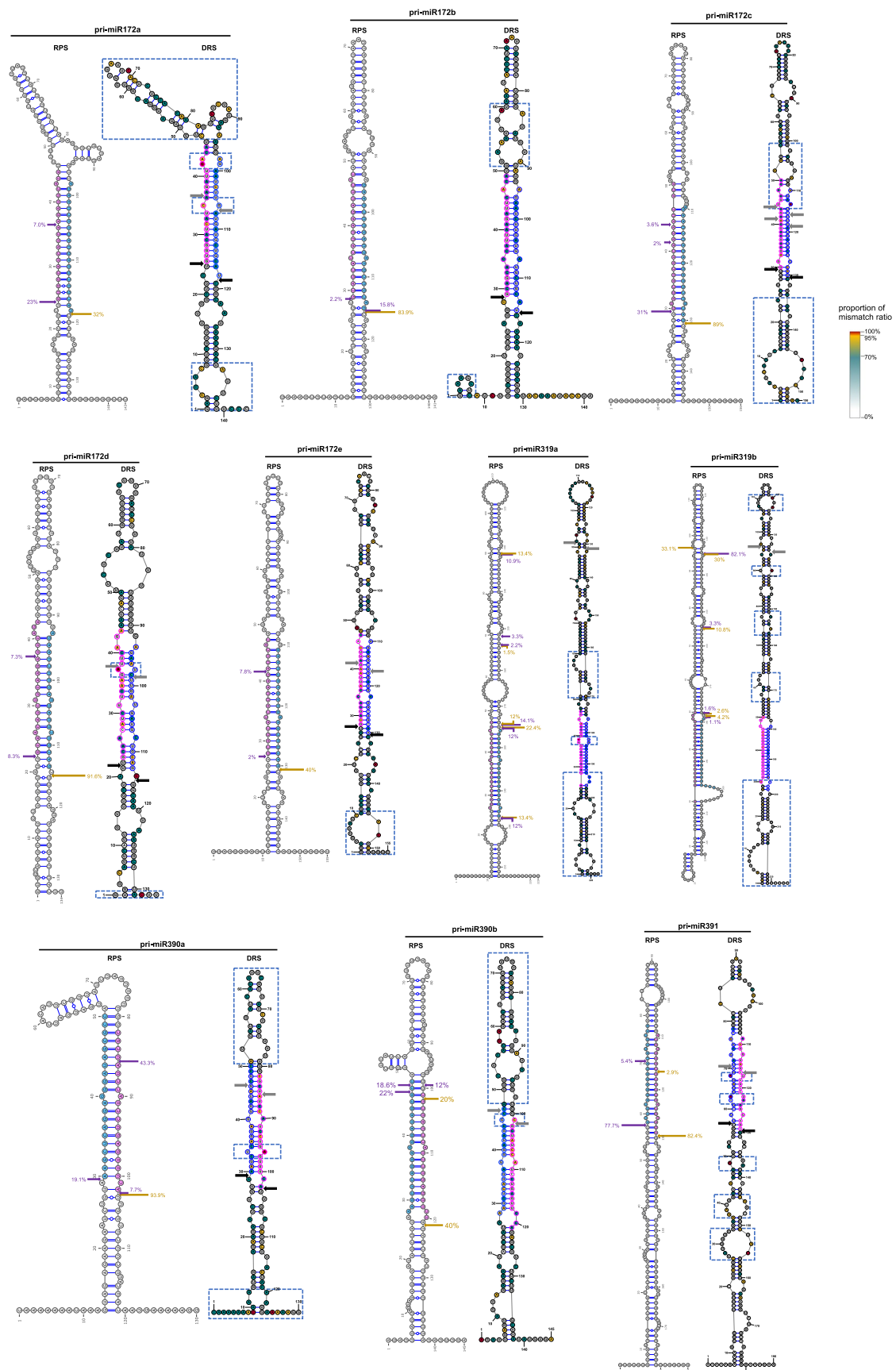
gray arrows indicated first cutting sites for BTL and LTB directions, respectively. RPS: RNAfold Predicted Structures. DRS: DMS Reactivity based Structures. Blue dotted boxes indicated structural differences in DRS.



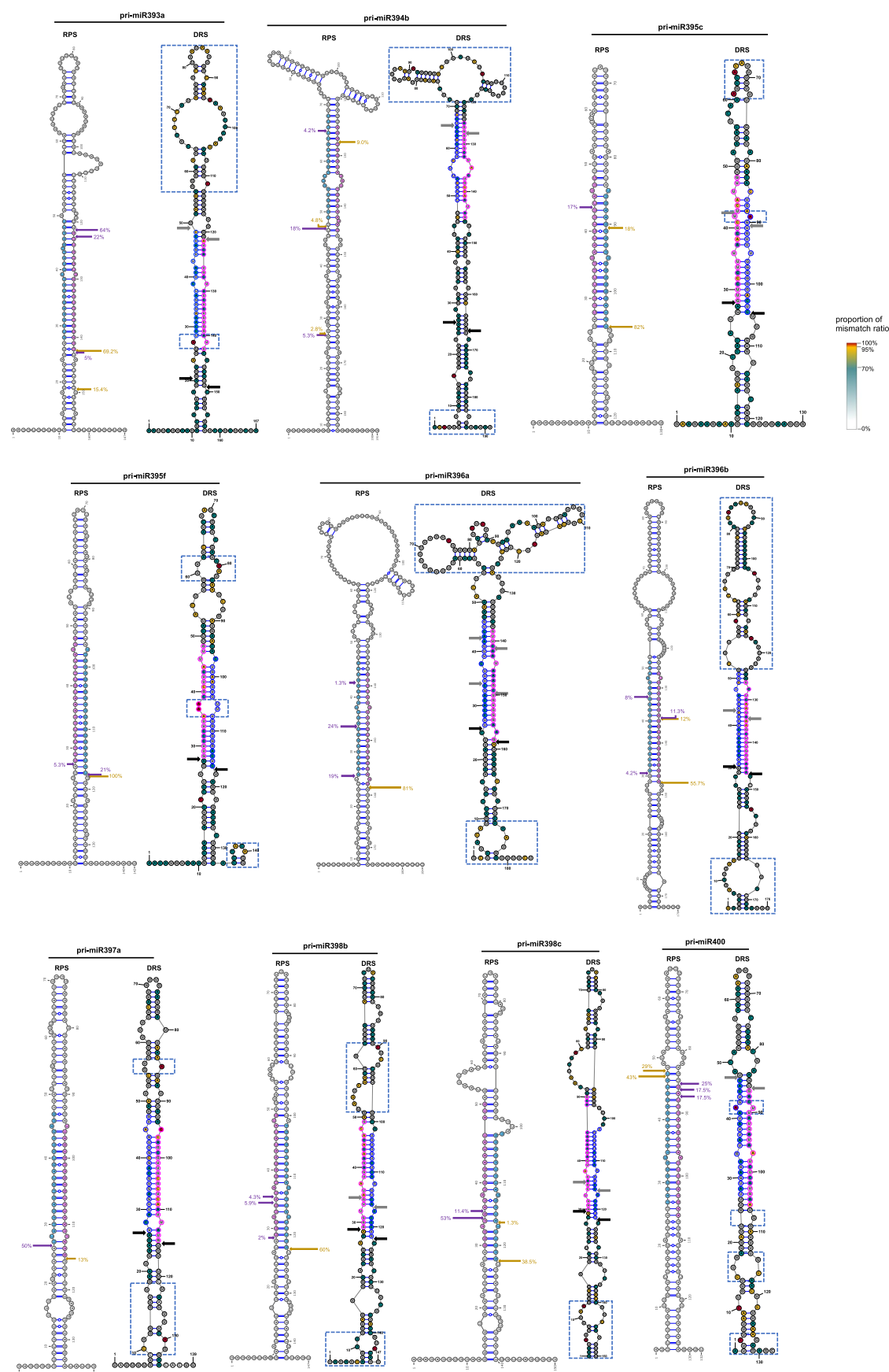
Extended Data Fig. 3 | Pri-miR165a, b, pri-miR166a, e, f, and pri-miR167a-d show structural differences in DRS (right) compared to RPS (left). Black and gray arrows indicated first cutting sites for BTL and LTB directions, respectively. RPS: RNAfold Predicted Structures. DRS: DMS Reactivity based Structures. Blue dotted boxes indicated structural differences in DRS.



Extended Data Fig. 4 | Pri-miR168b, pri-miR169a, d, and pri-miR171a-c show structural differences in DRS (right) compared to RPS (left). Black and gray arrows indicated first cutting sites for BTL and LTB directions, respectively. RPS: RNAfold Predicted Structures. DRS: DMS Reactivity based Structures. Blue dotted boxes indicated structural differences in DRS.

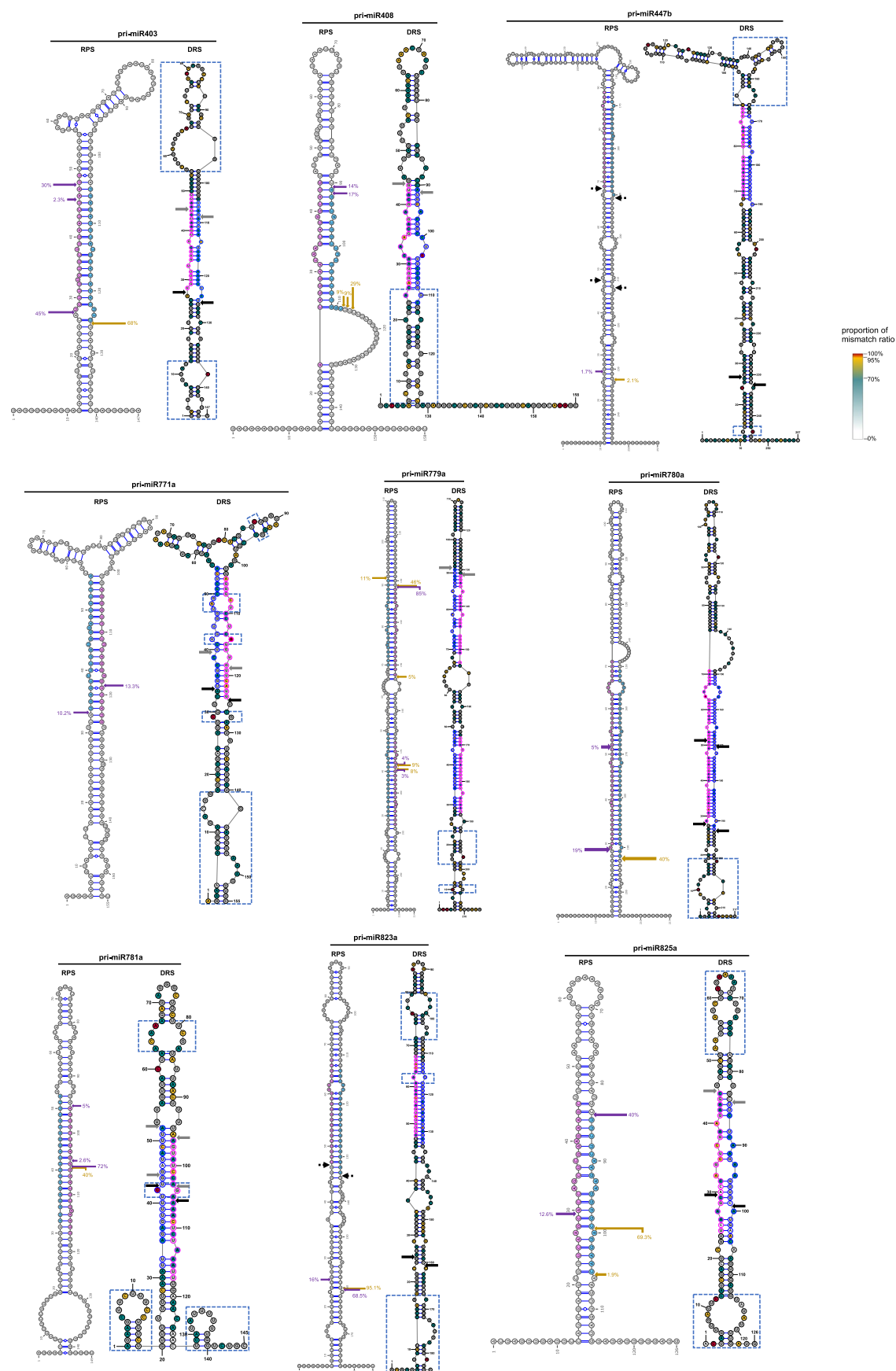


Extended Data Fig. 5 | Pri-miR172a-e, pri-miR319a, b, pri-miR390a, b and pri-miR391 show structural differences in DRS (right) compared to RPS (left). Black and gray arrows indicated first cutting sites for BTL and LTB directions, respectively. RPS: RNAfold Predicted Structures. DRS: DMS Reactivity based Structures. Blue dotted boxes indicated structural differences in DRS.



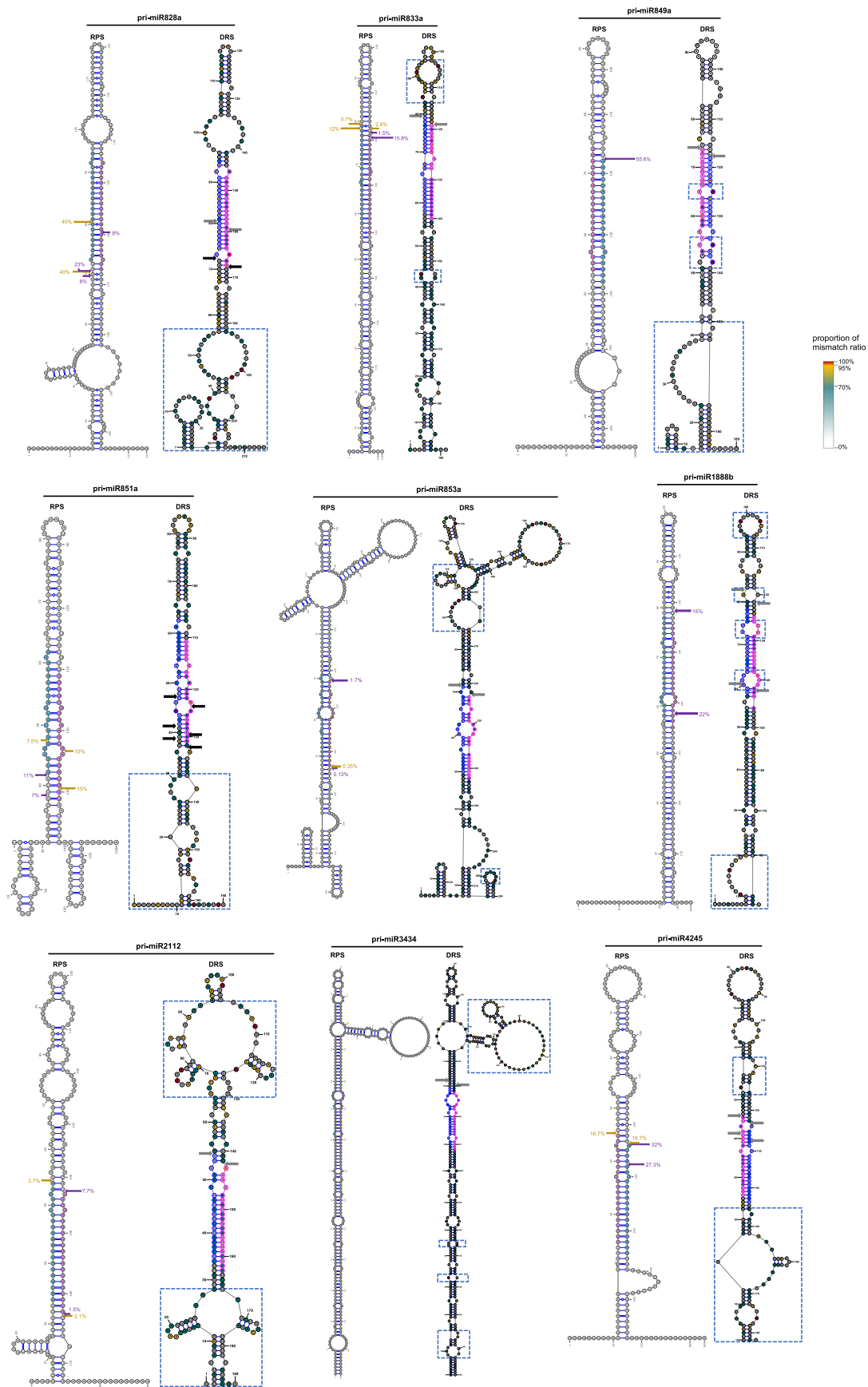
Extended Data Fig. 6 | Pri-miR393a, pri-miR394b, pri-miR395c, f, pri-miR396a, b, pri-miR397a, pri-miR398b, c and pri-miR400 show structural differences in DRS (right) compared to RPS (left). Black and gray arrows

indicated first cutting sites for BTL and LTB directions, respectively. RPS: RNAfold Predicted Structures. DRS: DMS Reactivity based Structures. Blue dotted boxes indicated structural differences in DRS.



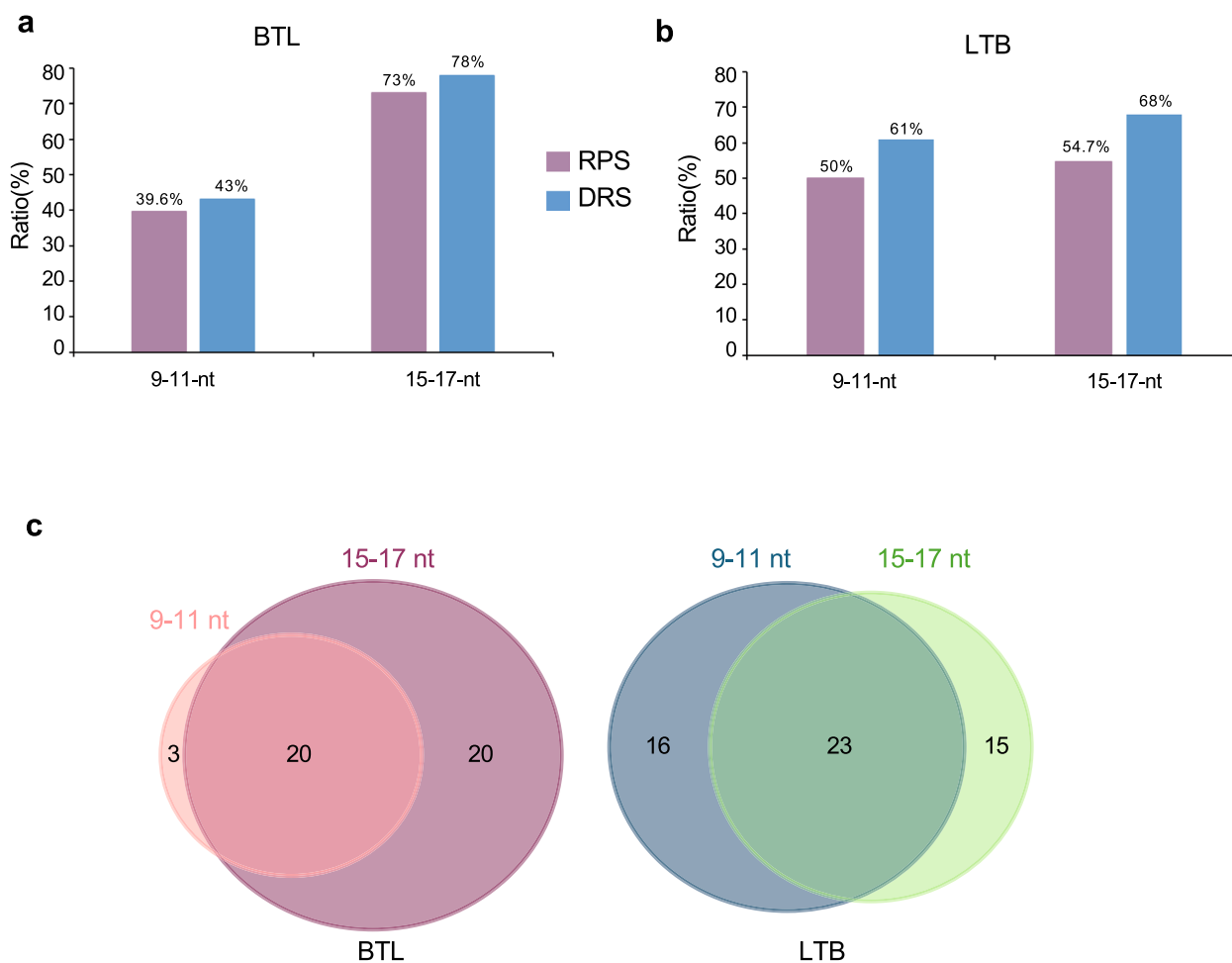
Extended Data Fig. 7 | Pri-miR403, pri-miR408, pri-miR447b, pri-miR771a, pri-miR779a, pri-miR780a, pri-miR781a, pri-miR823a and pri-miR825a show structural differences in DRS (right) compared to RPS (left). Black and gray

arrows indicated first cutting sites for BTL and LTB directions, respectively. RPS: RNAfold Predicted Structures. DRS: DMS Reactivity based Structures. Blue dotted boxes indicated structural differences in DRS.



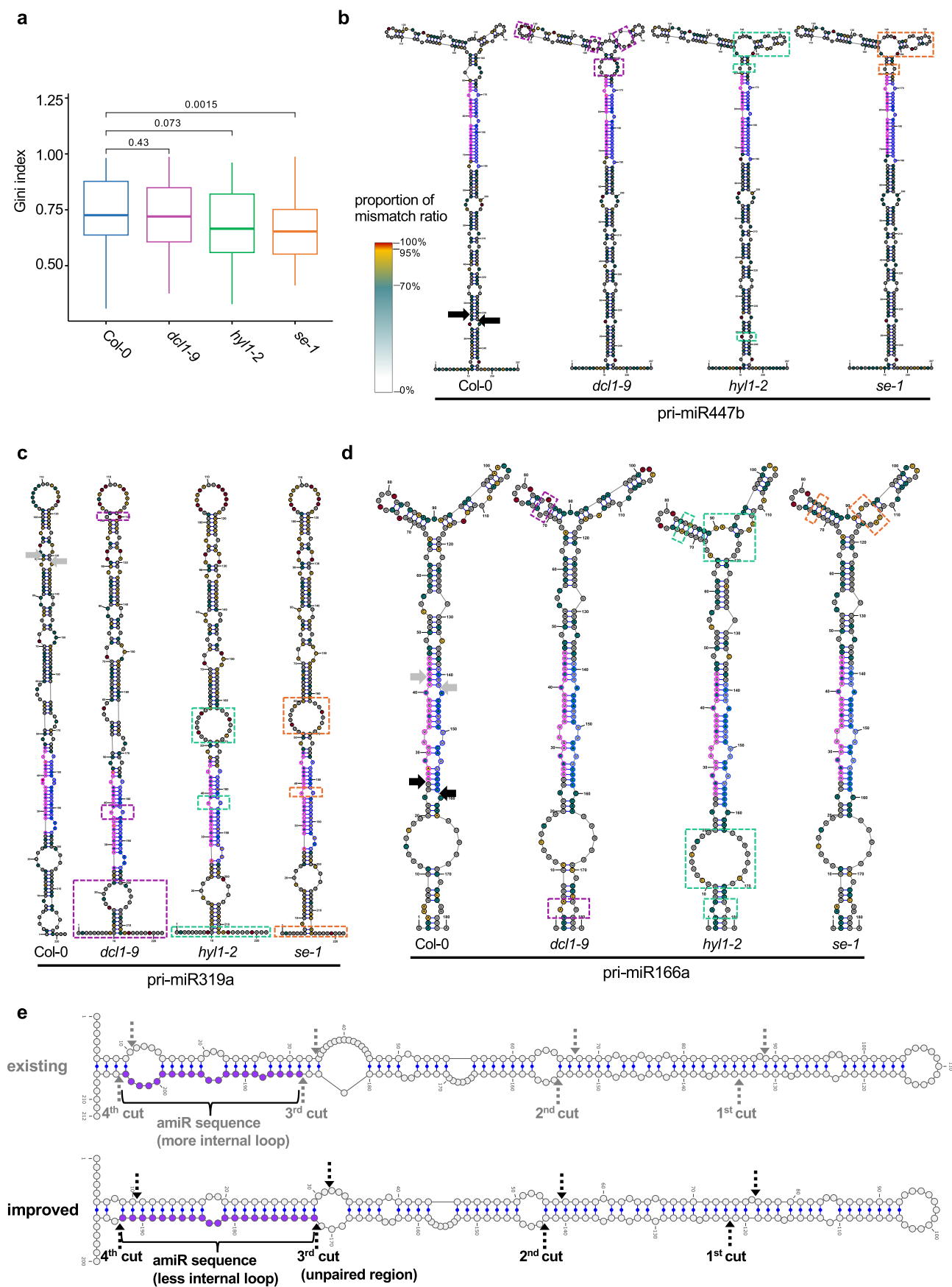
Extended Data Fig. 8 | Pri-miR828a, pri-miR833a, pri-miR849a, pri-miR851a, pri-miR853a, pri-miR1888b, pri-miR2112, pri-miR3434 and pri-miR4245 show structural differences in DRS (right) compared to RPS (left). Black and

gray arrows indicated first cutting sites for BTL and LTB directions, respectively. RPS: RNAfold Predicted Structures. DRS: DMS Reactivity based Structures. Blue dotted boxes indicated structural differences in DRS.



Extended Data Fig. 9 | *In vivo* RSS of pri-miRNAs can better explain the first cleavage sites than *in silico* predicted structures. (a) Barchart shows around 5% additional BTL-typed pri-miRNAs have internal loops/bulges that are 9-11 nt and 15-17 nt away from the first cleavage sites obtained in DRS vs RPS. (b) Barchart shows around 13% additional LTB-typed pri-miRNAs have internal loops/bulges

that are 9-11 nt and 15-17 nt away from the first cleavage sites obtained in DRS vs RPS. (c) Venn diagram shows that both BTL- and LTB-typed pri-miRNAs concurrently present internal loops/bulges that are -9-11 nt and -15-17 nt away from the first cutting sites. RPS: RNAfold Predicted Structures. DRS: DMS Reactivity based Structures.



Extended Data Fig. 10 | See next page for caption.

Extended Data Fig. 10 | DCL1, SE and HYL1 impact RSS of pri-miRNAs.

(a) Gini index of 67 common pri-miRNAs in Col-0, *dcl1-9*, *hyl1-2* and *se-1*. P value by Wilcoxon test. The data are from three biological replicates for Col-0, *dcl1-9* and *hyl1-2*, but two biological replicates for *se-1*. $P(dcl1-9 \text{ vs Col-0}) = 0.43$, $P(hyl1-2 \text{ vs Col-0}) = 0.073$, $P(se-1 \text{ vs Col-0}) = 0.0015$. P value by Wilcoxon test. Centres of the boxes represent the median values. Upper bound and lower bound show the first and the third quartiles respectively. Whiskers indicate data within

1.5× the interquartile range of both quartiles. **(b-d)** Examples of SBT1-processed pri-miR447b **(b)**, SLTB-processed pri-miR319a **(c)** and bidirectional-processed pri-miR166a **(d)** that show structural difference of pri-miRNAs in *dcl1-9*, *hyl1-2* and *se-1* compared to Col-0. Dotted boxes indicated structural differences in mutants. **(e)** Re-design of a known amiR backbone from pri-miR159a. An existing amiR backbone of pri-miR159a (top panel). Re-designing of the amiR backbone of pri-miR159a (bottom panel). amiR sequence is labelled with purple.

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (<i>n</i>) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided <i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>
☒	☐ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. <i>F</i> , <i>t</i> , <i>r</i>) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted <i>Give P values as exact values whenever suitable.</i>
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's <i>d</i> , Pearson's <i>r</i>), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection	Libraries for degradome-seq and DMS-MaPseq were sequenced on the Illumina system with pair-end 150 bp read length (NovaSeq PE150).
-----------------	--

Data analysis

Cutadapt (version 1.18)
 hisat2 (version 2.2.0)
 featureCounts (version 2.0.0)
 RNAfold web server (<http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi>)
 VARNA v3-93 (<https://varna.lisn.upsaclay.fr>)
 IGV (v2.6.3)
 psRNATarget V2 (<https://github.com/jtremblay/mirnatarget>)
 Fastx-toolkit (version 0.0.14) (Fastx-toolkit. http://hannonlab.cshl.edu/fastx_toolkit.)
 tophat2 (version 2.1.2)
 CountMismatch2Bed.py (https://github.com/changhaoli/TAMU_O2RSS)
 R package “ggpubr” (ggpubr: “ggplot2” based publication ready plots. <https://cran.r-project.org/web/packages/ggpubr/index.html>)
 RNAstructure (version 6.3)
 image J (v1.53t; <https://imagej.nih.gov/i>)
 PyMOL (version 2.5.7)
 SnapGene (4.3.6)
 Vector NTI Advance (version 11.5.1)
 BioRender (<https://biorender.com/>)
 Adobe illustrator CC (version 23.0.2)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The degradome-seq and DMS-MaPseq data were deposited in the NCBI Sequence Read Archive (SRA) under BioProject database with accession code PRJNA1092576. The Arabidopsis genome reference was obtained from TAIR (<https://www.arabidopsis.org>) and (CP002684–CP002688). Early annotated 326 pri-miRNAs information was from the miRBase (<https://www.mirbase.org/browse/results/?organism=ath>). sRNA-seq and AGOs-IP small RNA-seq data were obtained from NCBI website with GEO accession GSE78090, GSE66599 and GSM707678–GSM707691. All other data supporting the findings of the study are present in the main text and/or the Supplementary Information.

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender N/A

Reporting on race, ethnicity, or other socially relevant groupings N/A

Population characteristics N/A

Recruitment N/A

Ethics oversight N/A

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size For transgenic plants, five independent transgenic lines were characterized for each semi-active DCL1 point mutants. sRNA blot experiments

Sample size	were conducted with three biologic replicates. Two biologic replicates for degradome sequencing, three biologic replicates for DMS-MaPseq. Box plots were illustrated, showcasing five values arranged from top to bottom: maximum, third quartile, median, first quartile, and minimum. The lower whisker represents the minimum value, while the upper whisker denotes the maximum value, potentially including outliers—data points identified as statistically deviating beyond the upper and lower whiskers.
Data exclusions	No data exclusion in this study.
Replication	All analyses in this study were conducted with two or three biological replicates, as specified in the respective figure legends. All attempts at replication were successful.
Randomization	For library construction, plants in the indicated background were grown in one tray containing 32 or 36 plants, and the plants in the same tray is one replicate.
Blinding	Blinding wasn't implemented during experiments and data analysis, but the investigators were blinded to group allocation while collecting samples.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☐	☒ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Anti-Myc antibody (rabbit Ab, Cat# C3956, Sigma-Aldrich, 1:10,000 dilution, clone name: not provided by the manufacturer). Anti-actin antibody (mouse Ab, Cat# A0480, Sigma-Aldrich, 1:10,000 dilution, clone name: 10-B3 (MAbGPa)). Anti-rabbit Ig G, Horseradish peroxidase linked whole antibody (from donkey) (Cat#: NA934v, Cytiva, 1:5,000 dilution, clone name: not provided by the manufacturer). Anti-mouse Ig G, Horseradish peroxidase linked whole antibody (from sheep) (Cat#: NA931v, Cytiva, 1:5,000 dilution, clone name: not provided by the manufacturer).
Validation	All antibodies were validated in Western Blot in Arabidopsis. Their validation statements and experiment tests can also be found in the manufacturers' websites. Anti-Myc (Sigma, Cat# C3956). The antibody is validated at https://www.sigmaaldrich.com/US/en/product/sigma/c3956 . Anti-actin (Sigma, Cat# A0480). The antibody is validated at https://www.sigmaaldrich.com/US/en/product/sigma/a0480 . Anti-rabbit (Cytiva, Cat# NA934v) and Anti-mouse (Cytiva, Cat# NA931v). The antibody is validated at https://www.cytivalifesciences.com/en/us/shop/protein-analysis/blotting-and-detection/blotting-standards-and-reagents/amersham-ecl-hrp-conjugated-antibodies-p-06260 .

Plants

Seed stocks	Arabidopsis thaliana (A. thaliana) ecotype Columbia (Col-0); dcl1-9 (CS3828); se-1(CS3257); se-2 (SAIL_44_G12); hyl1-2 (SALK_064863)
Novel plant genotypes	Novel transgenic lines including pBA002a-PDCL1-FM-gDCL1 (E1507Q) and pBA002a-PDCL1-FM-gDCL1 (E1696Q) were transformed into the dcl1-9(+/-) background by the floral dip transformation method, each one get 5 independently transgenic lines.
Authentication	For all seed stocks, they were used for this study as verified previously(Wang, L. et al. PRP4KA phosphorylates SERRATE for degradation via 20 S proteasome to fine- tune miRNA production in Arabidopsis). For transgenic plants, they were screened by western blot with an anti-Myc antibody (Sigma, Cat# C3956) and PCR analysis with primers listed in Supplementary Table6.

Parallel degradome-seq and DMS-MaPseq substantially revise the miRNA biogenesis atlas in *Arabidopsis*

In the format provided by the
authors and unedited

This file includes:

Supplementary Figures

Supplementary Figure 1. Quality control analysis shows that degradome sequencing libraries have high quality and repeatability.

Supplementary Figure 2. Degradome sequencing reveals that 37 pri-miRNAs have BTL processing pattern (Be noted that this figure contains five pages).

Supplementary Figure 3. Degradome sequencing reveals that 16 pri-miRNAs have SBTL processing pattern (Be noted that this figure contains two pages).

Supplementary Figure 4. Degradome sequencing reveals that 38 pri-miRNAs have LTB processing pattern (Be noted that this figure contains five pages).

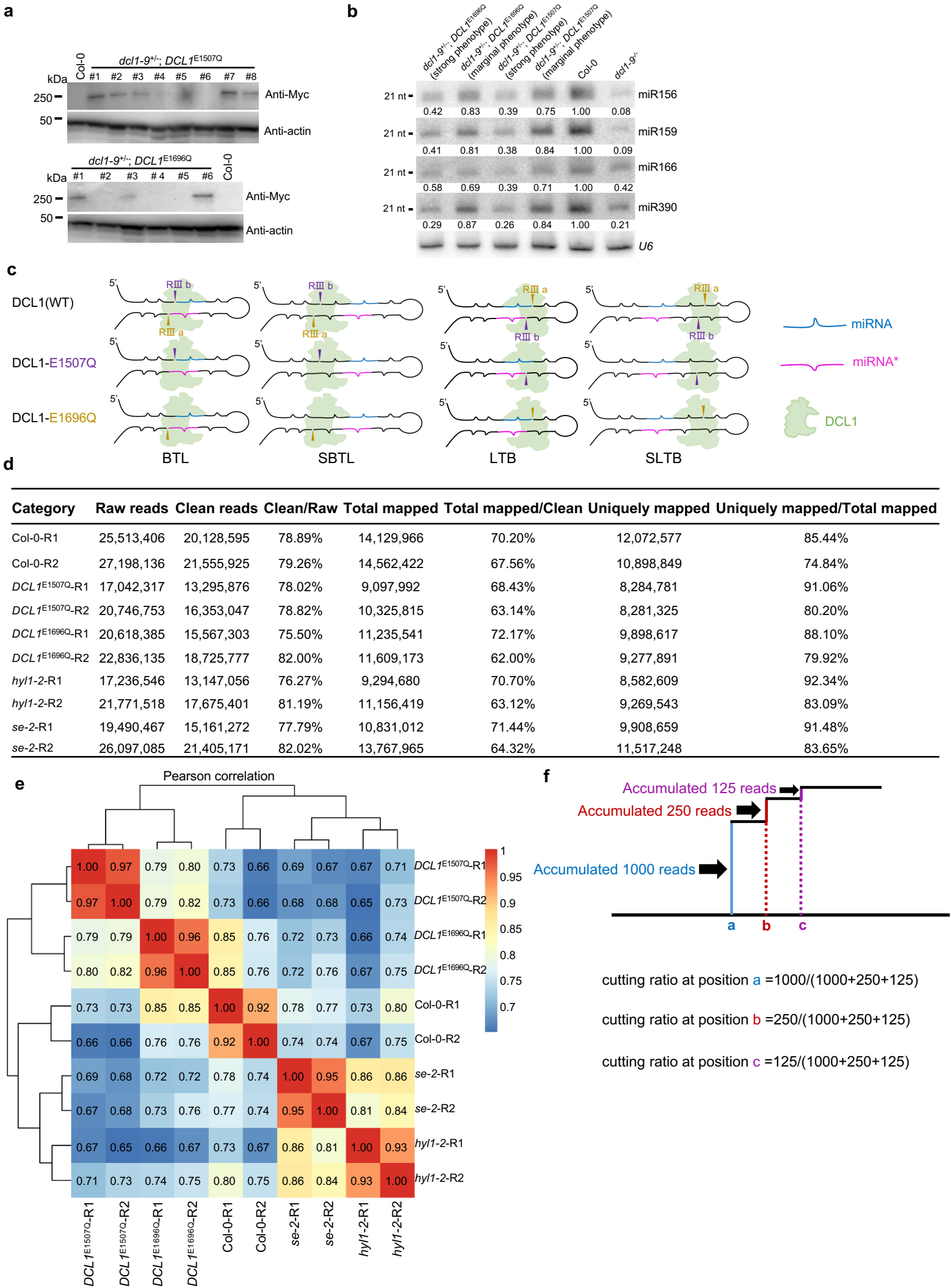
Supplementary Figure 5. Degradome sequencing reveals that 13 pri-miRNAs have the SLTB processing pattern (Be noted that this figure contains two pages).

Supplementary Figure 6. Degradome sequencing reveals that 43 pri-miRNAs have bidirectional processing pattern and pri-miR867 possibly yields new miRNA clusters (Be noted that this figure contains eight pages).

Supplementary Figure 7. SE and HYL1 show different impacts on pri-miRNA processing.

Unprocessed western blots for Supplementary Fig. 1a

Supplementary Fig. 1



Supplementary Figure 1. Quality control analysis shows that degradome sequencing libraries have high quality and repeatability.

(a) Screening of elite lines expressing dominant-negative forms (*DCL1*^{E1507Q} and *DCL1*^{E1696Q}) via western blot analysis using an anti-Myc antibody. Col-0 is a negative control. Actin serves as a loading control.

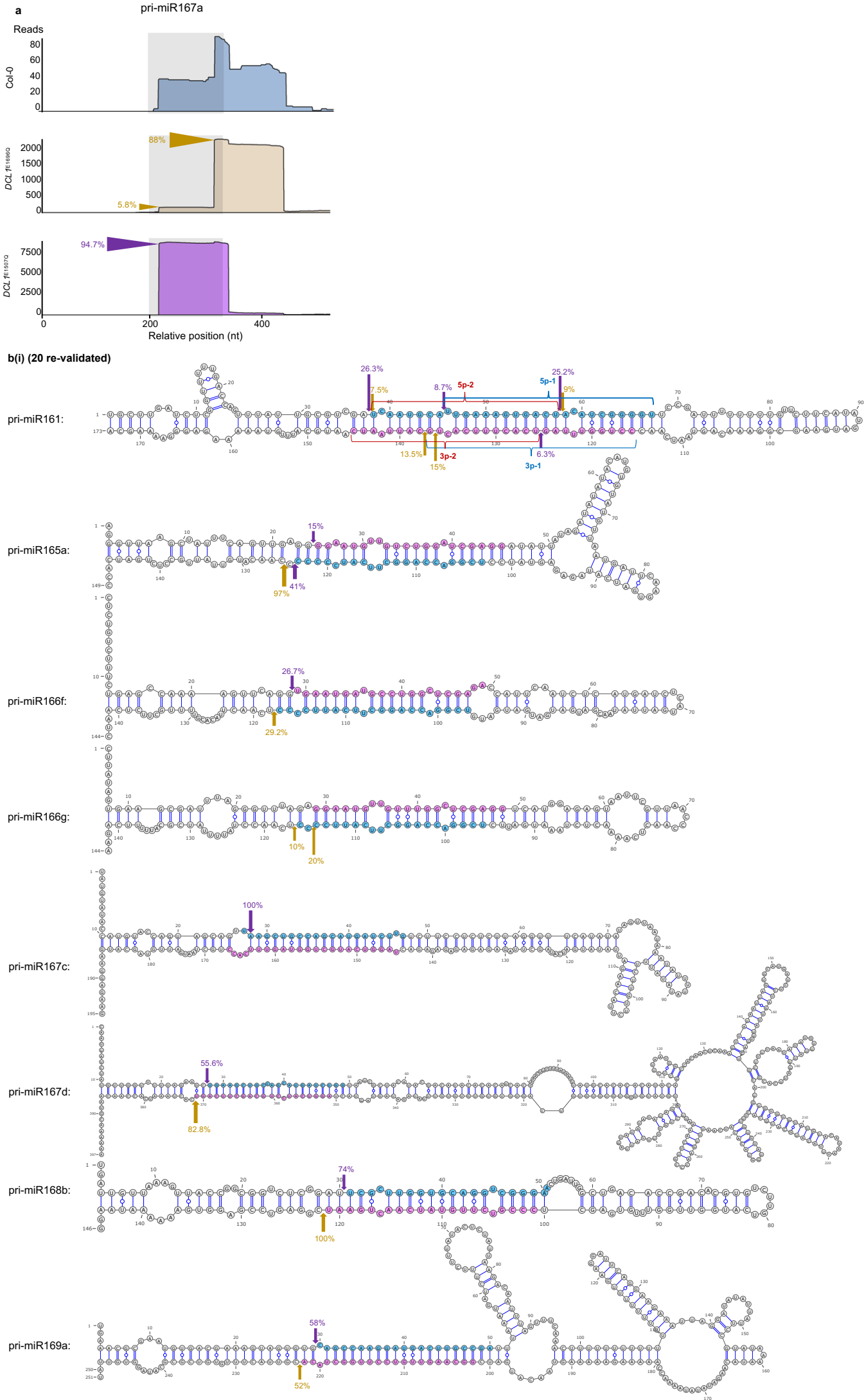
(b) sRNA blot analyses of the selected miRNAs in the indicated lines. *U6* is a loading control. nt, nucleotides. The experiment was independently repeated by three times with similar results.

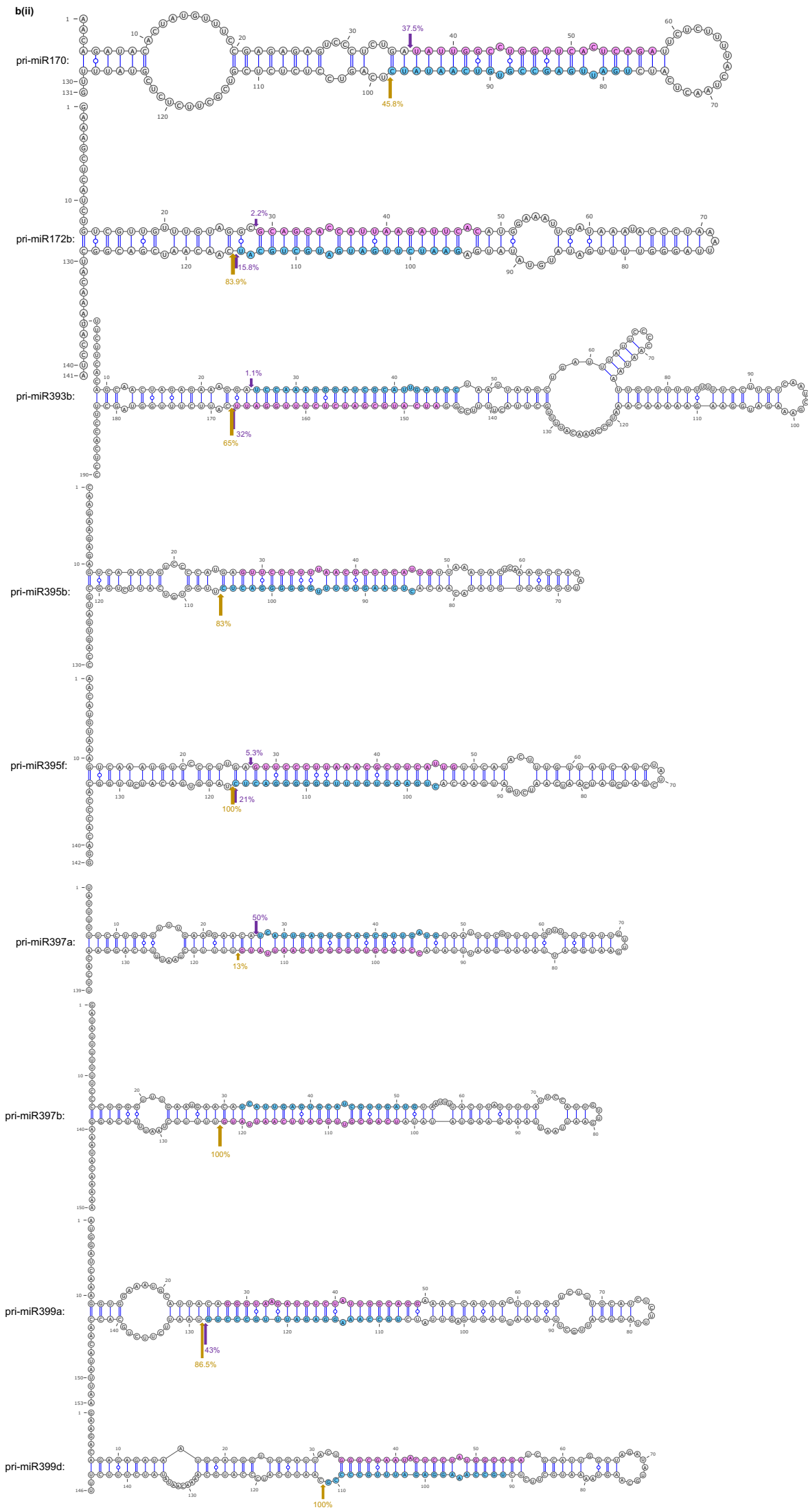
(c) Schematic illustration of the cleavage products by DCL1 (WT) (top), DCL1-E1507Q (middle) and DCL1-E1696Q (bottom). Yellow-brown and purple arrowheads show the cleavage sites by RNase IIIa and RNase IIIb, respectively. BTL: base to loop; SBTL: sequential base to loop; LTB: loop to base; SLTB: sequential loop to base.

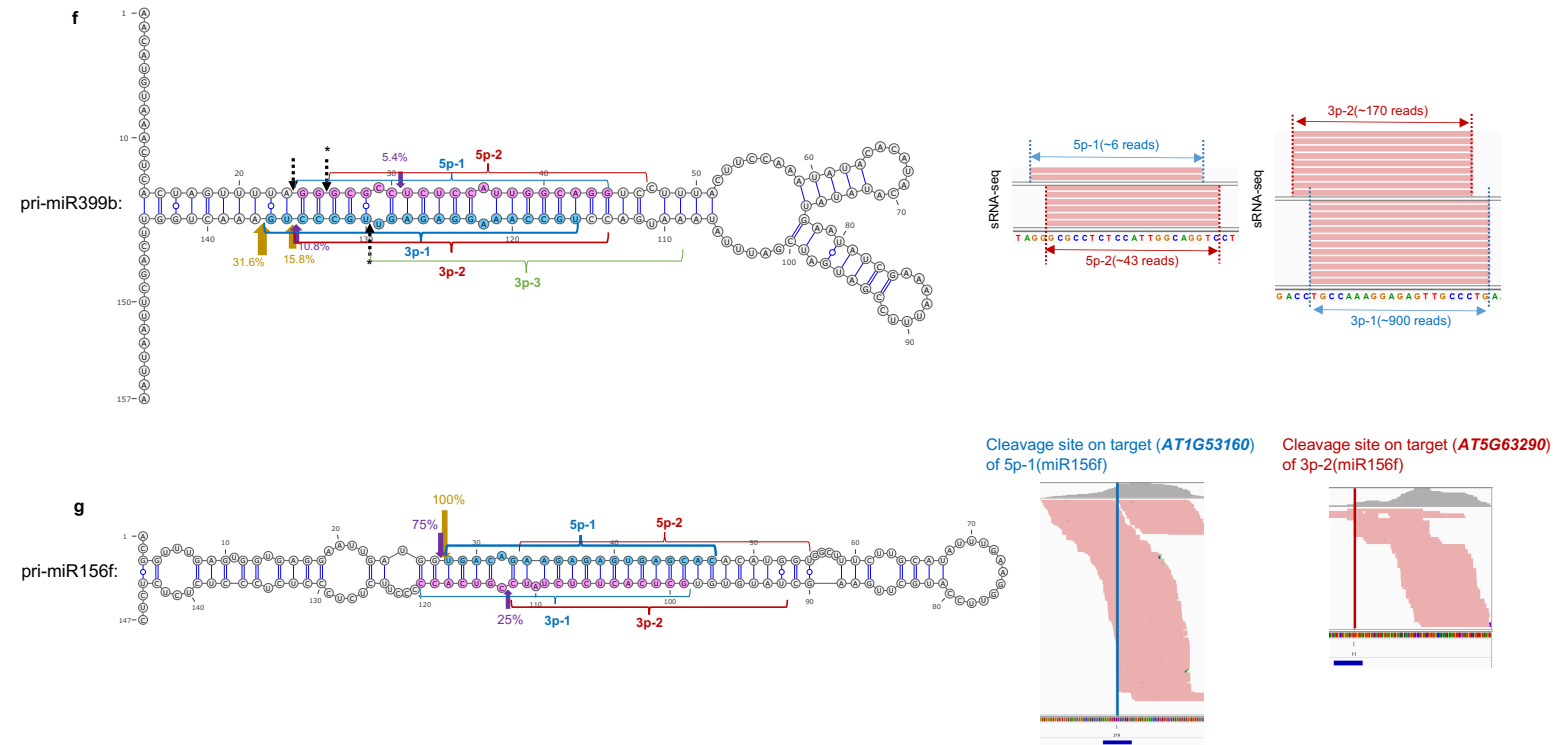
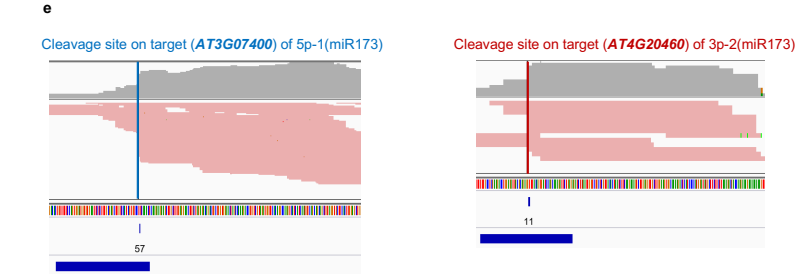
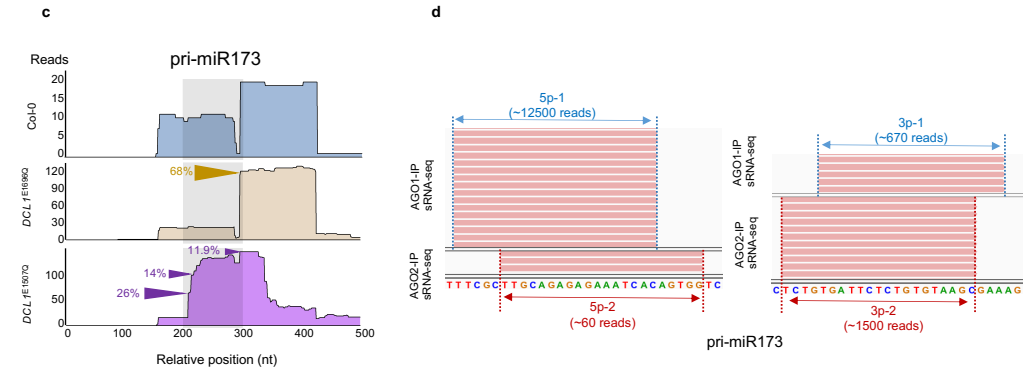
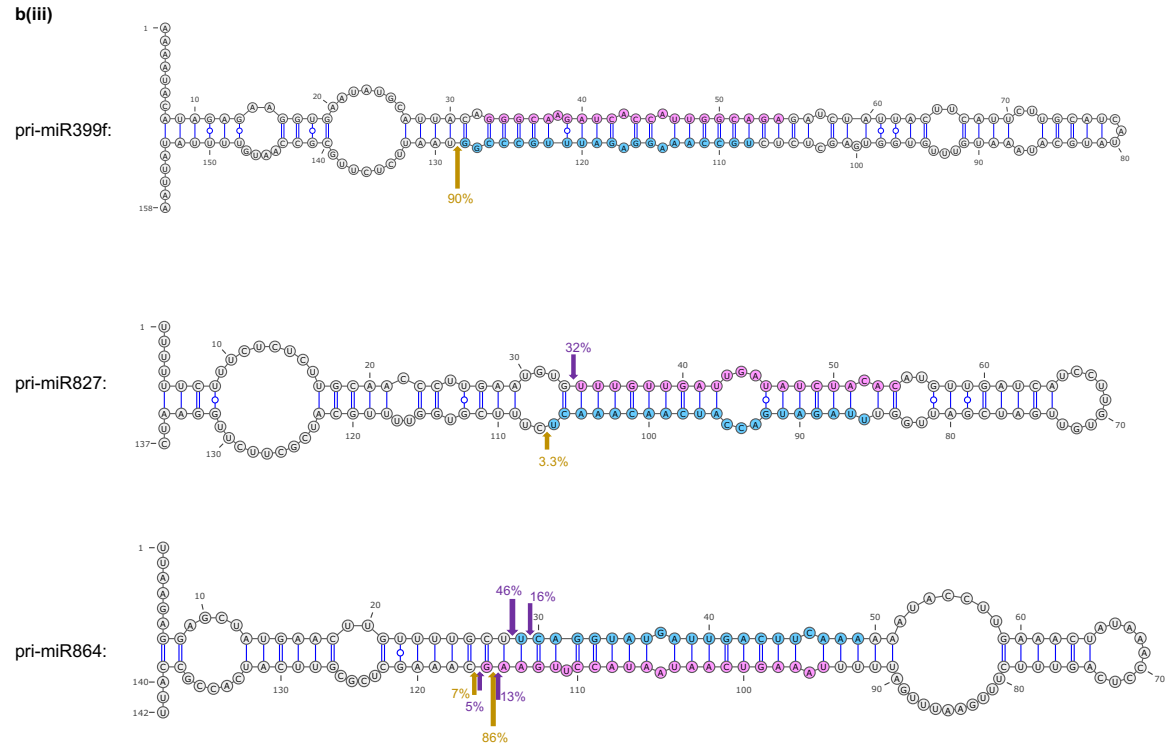
(d) Table shows total and unique reads that match the reference transcripts for all samples. R1: replicate 1. R2: replicate 2.

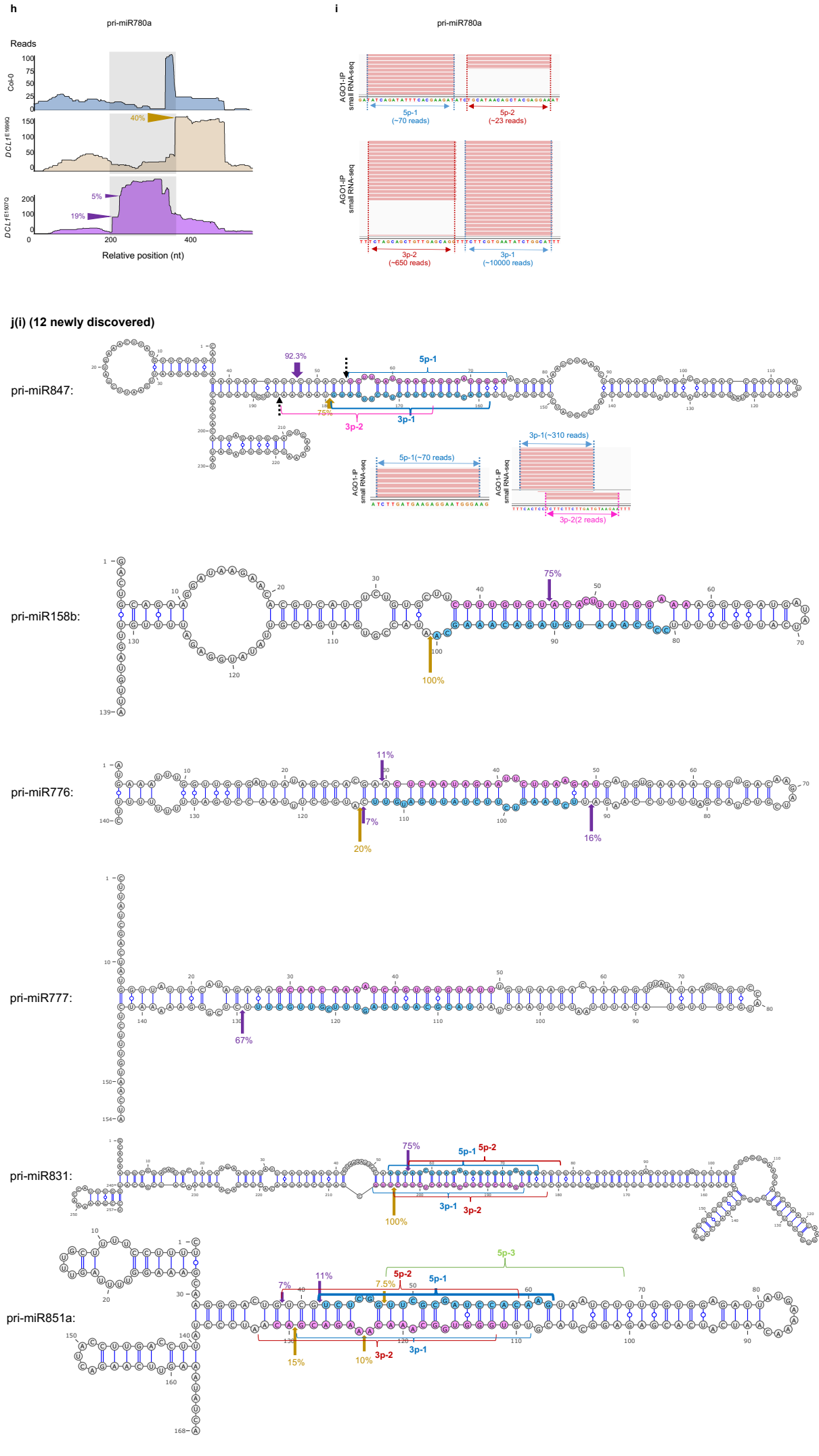
(e) Heatmap shows high Pearson correlation (≥ 0.92) between two biological replicates for all samples.

(f) Schematic illustration of algorithm for ratio calculation of the cleavage sites.

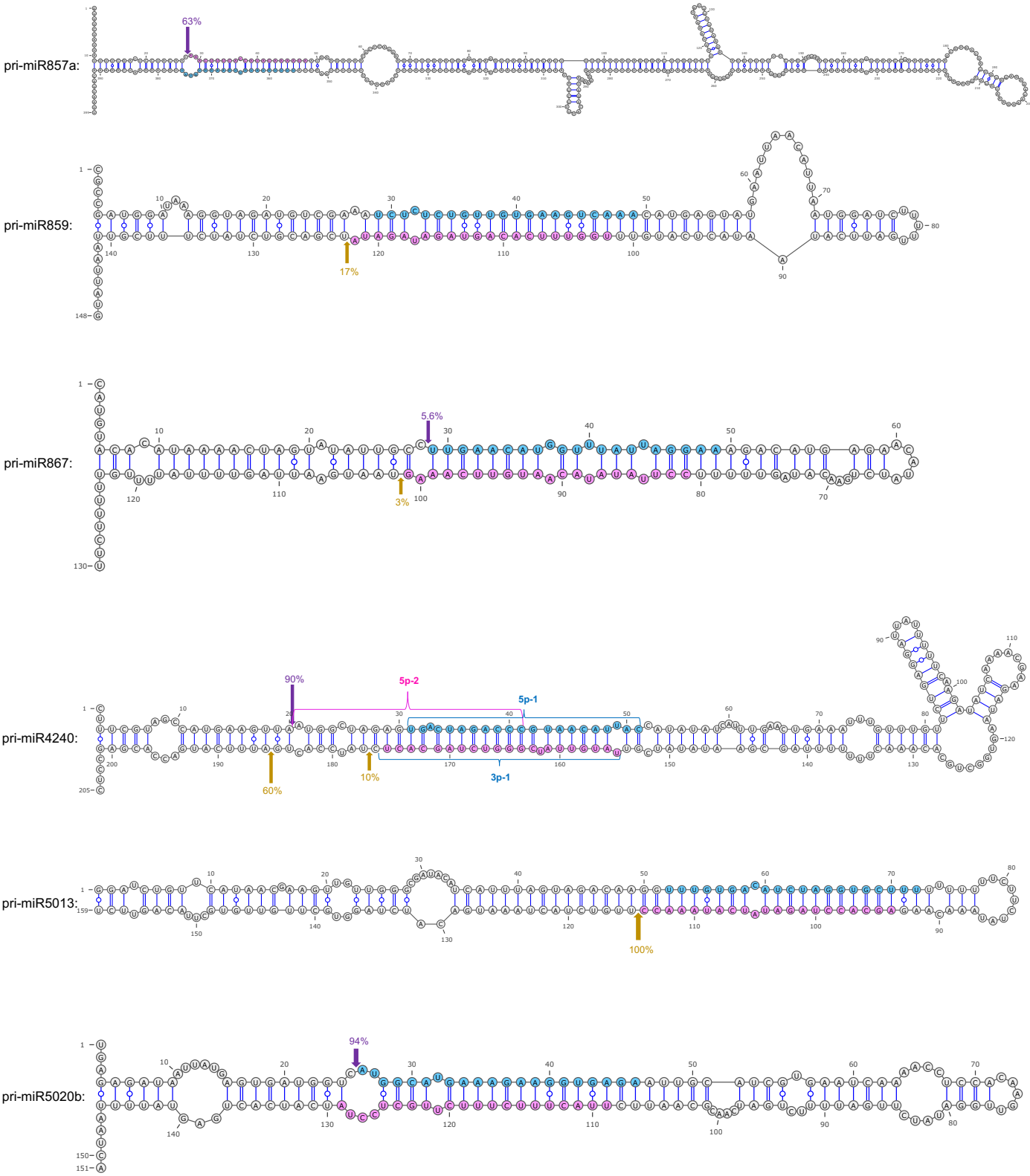








j(ii)



Supplementary Figure 2. Degradome sequencing reveals that 37 pri-miRNAs have BTL processing pattern.

(a) IGV graphs of reads' numbers and relative ratios at detected cleavage sites of pri-miR167a in Col-0 and semi-active DCL1 expressing lines. Related to Fig. 2a.

(b) A total of 20 BTL-processed pri-miRNAs discovered earlier are re-validated here. Be noted: there are three pages for this panel (2b(i) to 2b(iii)).

(c) IGV graphs of reads' numbers and relative ratios at detected cleavage sites of pri-miR173 in Col-0 and semi-active DCL1 expressing lines. Related to Fig. 2b.

(d) IGV graphs of two species of miRNA/*s from pri-miR173. The canonical miR173/* (5p-1/3p-1) is sorted into AGO1 whereas the non-canonical pair (5p-2/3p-2) newly discovered here is recovered from AGO2. Related to Fig. 2b.

(e) Degradome-seq data verified that the newly identified 3p-2 of pri-miR173 can target a different mRNA (*AT4G20460*) for cleavage compared to annotated 5p-1 (one representative target is *AT3G07400*). Blue rectangular boxes represent the target range, numbers represent read counts for each cleavage site. Related to Fig. 2b.

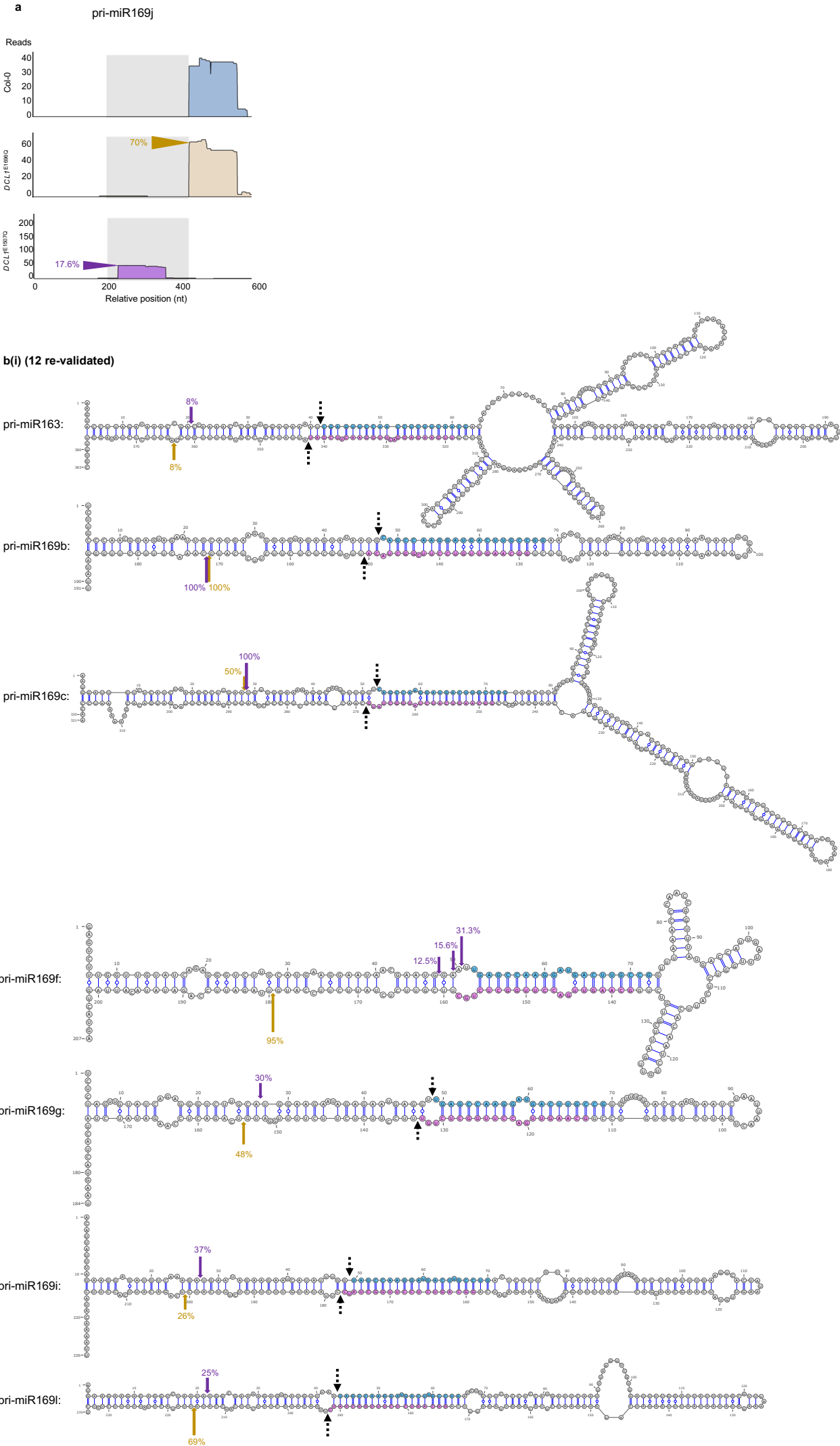
(f) Pri-miR399b is now re-annotated to three kinds of BTL processing patterns (left panel) and resultant two productive species of miRNA/*s shown in the IGV graphs (right panel). One functional miRNA/* is the canonical miR399/*, designated as the 5p-1 and 3p-1 products (blue brackets) while the other is a new set of miRNA/* duplex in red that could be recovered in sRNA-seq and that we referred as 5p-2 and 3p-2. The reads of the recovered miRNA/*s (5p-1/3p-1; 5p-2/3p-2) are shown on the right panels. Of note, pri-miR399b also underwent the third mode of processing that is shifted 7 nt toward the canonical miR399/* duplex, producing a 3p-3 abortive product in the 3' arm (light green brackets). The presence of multiple starting processing sites for pri-miR399b is likely due to its complicated terminal loop structure. The black dotted arrows denoted the intended cutting positions but were not identified in our data. The cleavage site marked with an asterisk (*) was also discovered in WT and *fiery1*¹⁴.

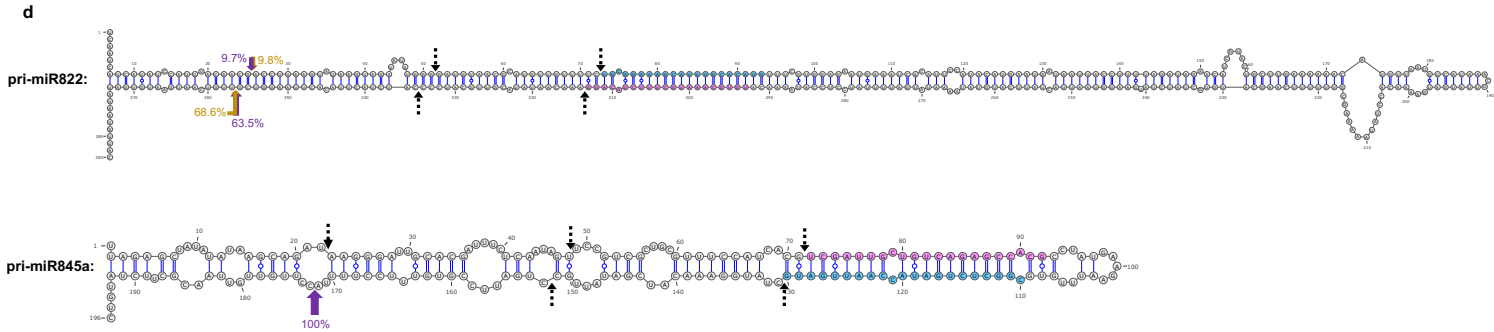
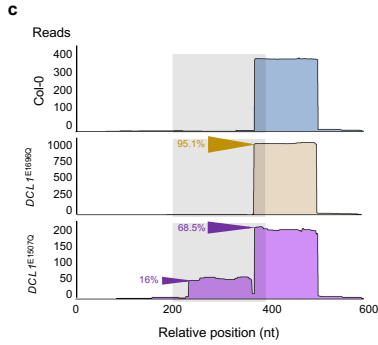
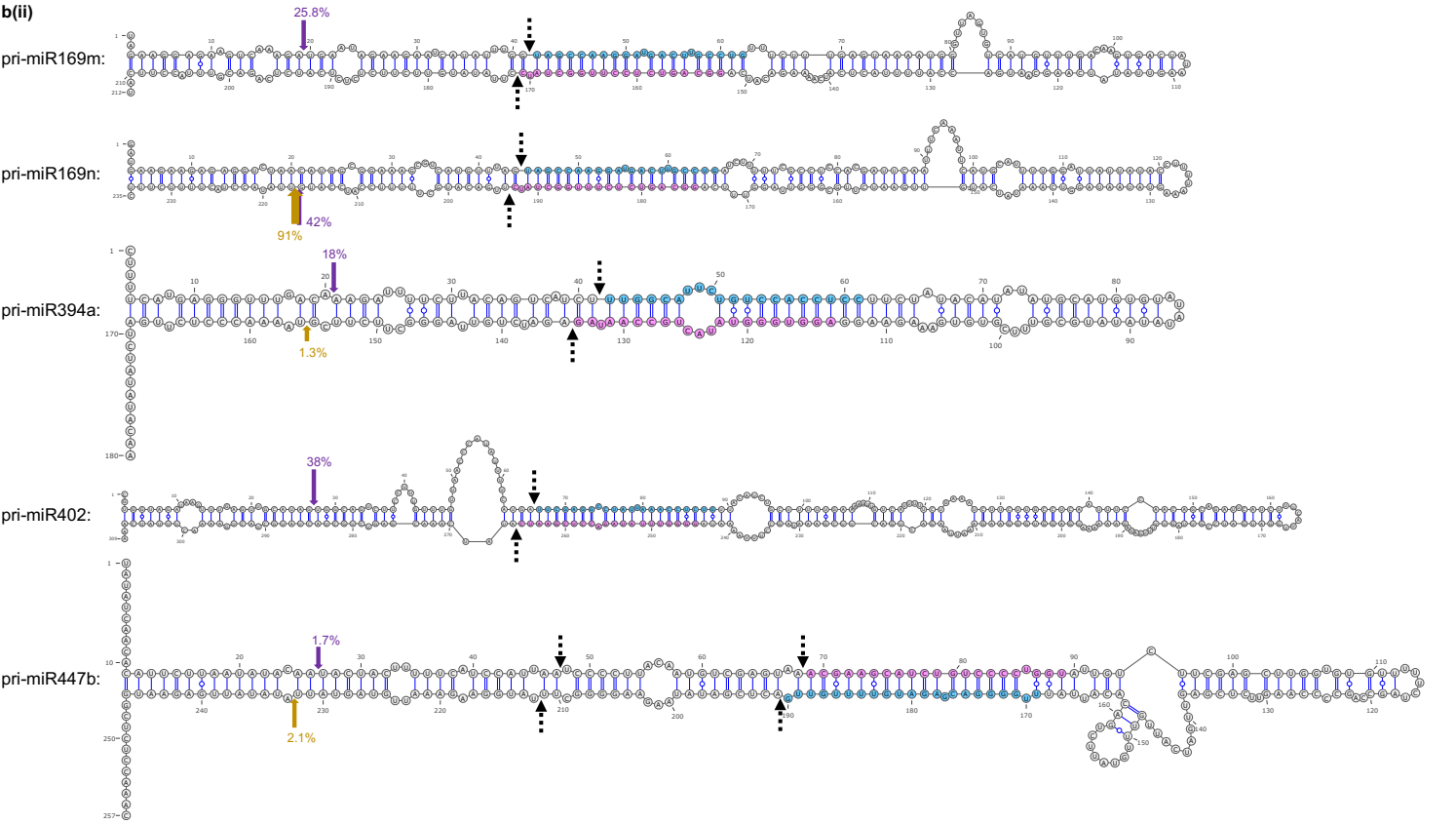
(g) Pri-miR156f is now re-annotated to the two types of BTL processing, resulting in miR156f/* (5p-1/3p-1) and a newly productive miRNA/* species (5p-2/3p-2) (left panel) that could be recovered from sRNA-seq. Be noted that 5p-1 targets mRNA (*AT1G53160*), 3p-2 cleaves mRNA (*AT5G63290*) according to the degradome-seq (right panel).

(h) IGV graphs of reads' numbers and relative ratios at detected cleavage sites of pri-miR780a in Col-0 and semi-active DCL1 expressing lines. Related to Fig. 2c.

(i) IGV graphs of two species of miRNA/*s from pri-miR780a, both (5p-1/3p-1) and (5p-2/3p-2) are sorted into AGO1. Related to Fig. 2c.

(j) Newly identified BTL patterns for 12 pri-miRNAs with some new targets included. Be noted that there are two pages for this panel (2j(i) and 2j(ii)). Pri-miR158b had an alternative starting cleavage site that is in the middle of miR158/*, likely resulting in an abortive product, which is not recovered in RISC. Pri-miR831 does not generate the products (5p-1/3p-1) as predicted from miRBase. Rather, the initial cleavage position shifted 3 nt away from the predicted site and generated a new pair but functional duplex (5p-2/3p-2, in red). Pri-miR847 follows two kinds of BTL modes, producing functional 5p-1/3p-1 recovered from AGO1-RISC (in blue), but abortive 3p-2 (in pink). The processing of pri-miR851a is even more heterogeneous as it has three different sites for initial cleavages with comparable ratios: one produces miR851a/* as predicted (5p-1/3p-1, in blue); another way shifted 3 nt toward the base region and generated a new pair but productive form of miRNA/* (enclosed by red brackets). Importantly, 5p-1 targets *AT4G21350* whereas the new 5p-2 cleaves *AT2G34750*. Additionally, the third processing shifted 6 nt onto the canonical miRNA/* duplex but yielded an abortive product (marked by a light green bracket). Finally, pri-miR4240 displays two ways of BTL processing: one produces the canonical miRNA/* (in blue), but the other generates abortive product with only one strand of sRNA that could be recovered in RISC but with low read counts (labelled in pink).





Supplementary Figure 3. Degradome sequencing reveals that 16 pri-miRNAs have SBTL processing pattern.

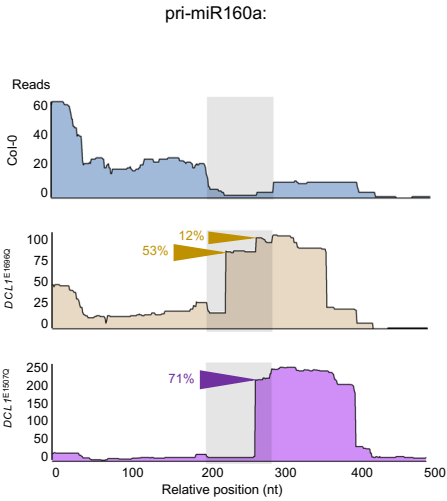
(a) IGV graphs of reads' numbers and relative ratios at detected cleavage sites of pri-miR169j in Col-0 and semi-active DCL1 expressing lines. Related to Fig. 2d.

(b) The rest 12 SBTL-processed pri-miRNAs discovered earlier are re-validated here. Be noted: there are two pages for this panel (3b(i) and 3b(ii)).

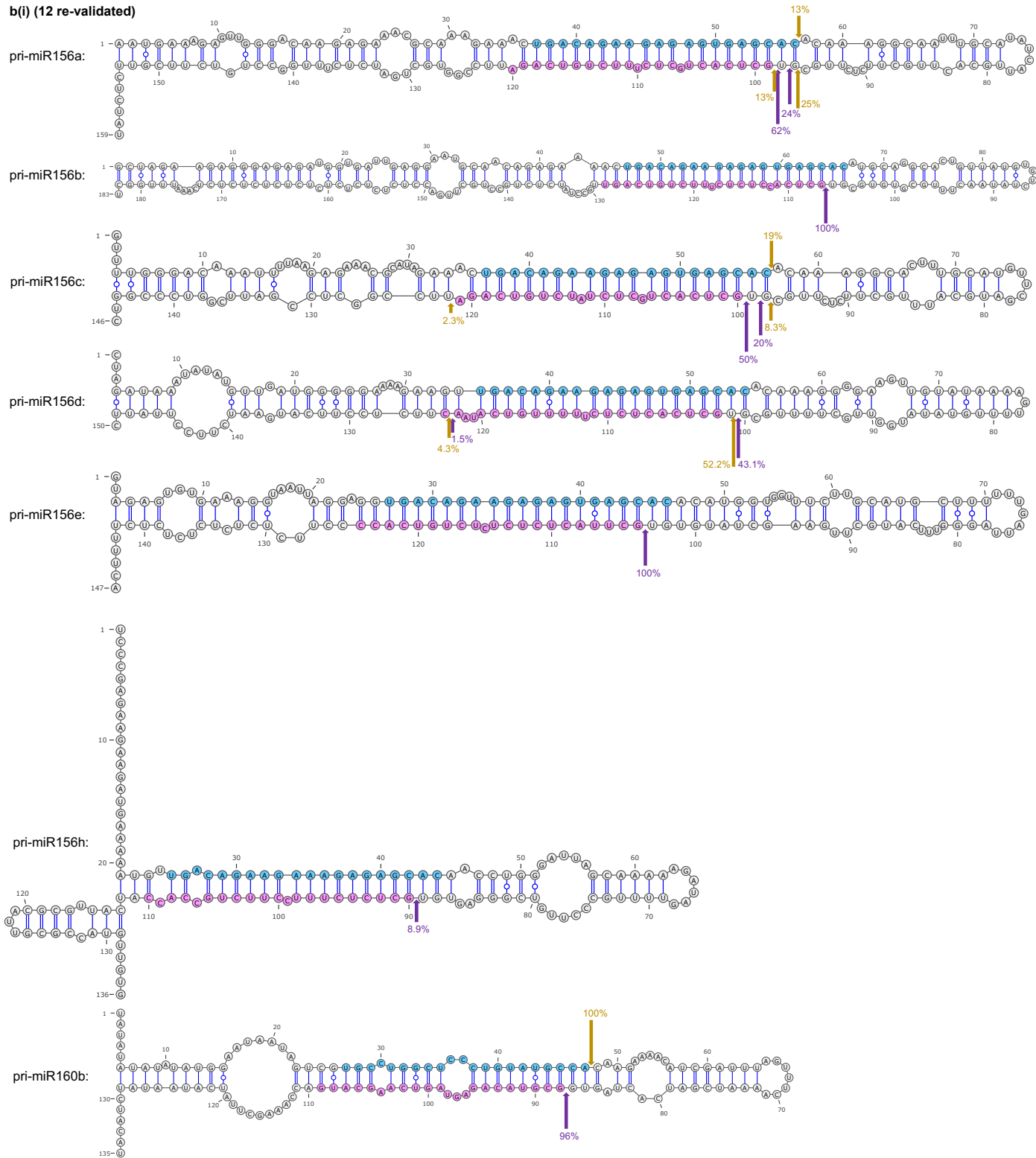
(c) IGV graphs of reads' numbers and relative ratios at detected cleavage sites of pri-miR823 in Col-0 and semi-active DCL1 expressing lines. Related to Fig. 2e.

(d) Two more newly identified examples following SBTL processing pattern, including pri-miR822 (top) and pri-miR845a (bottom).

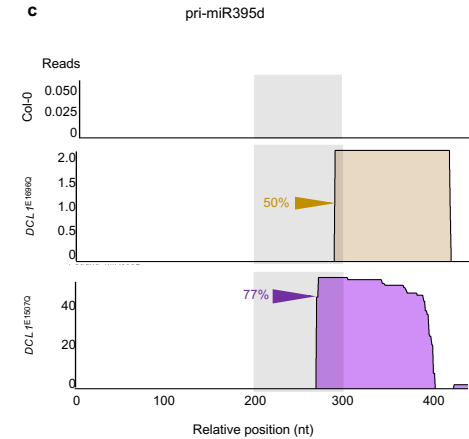
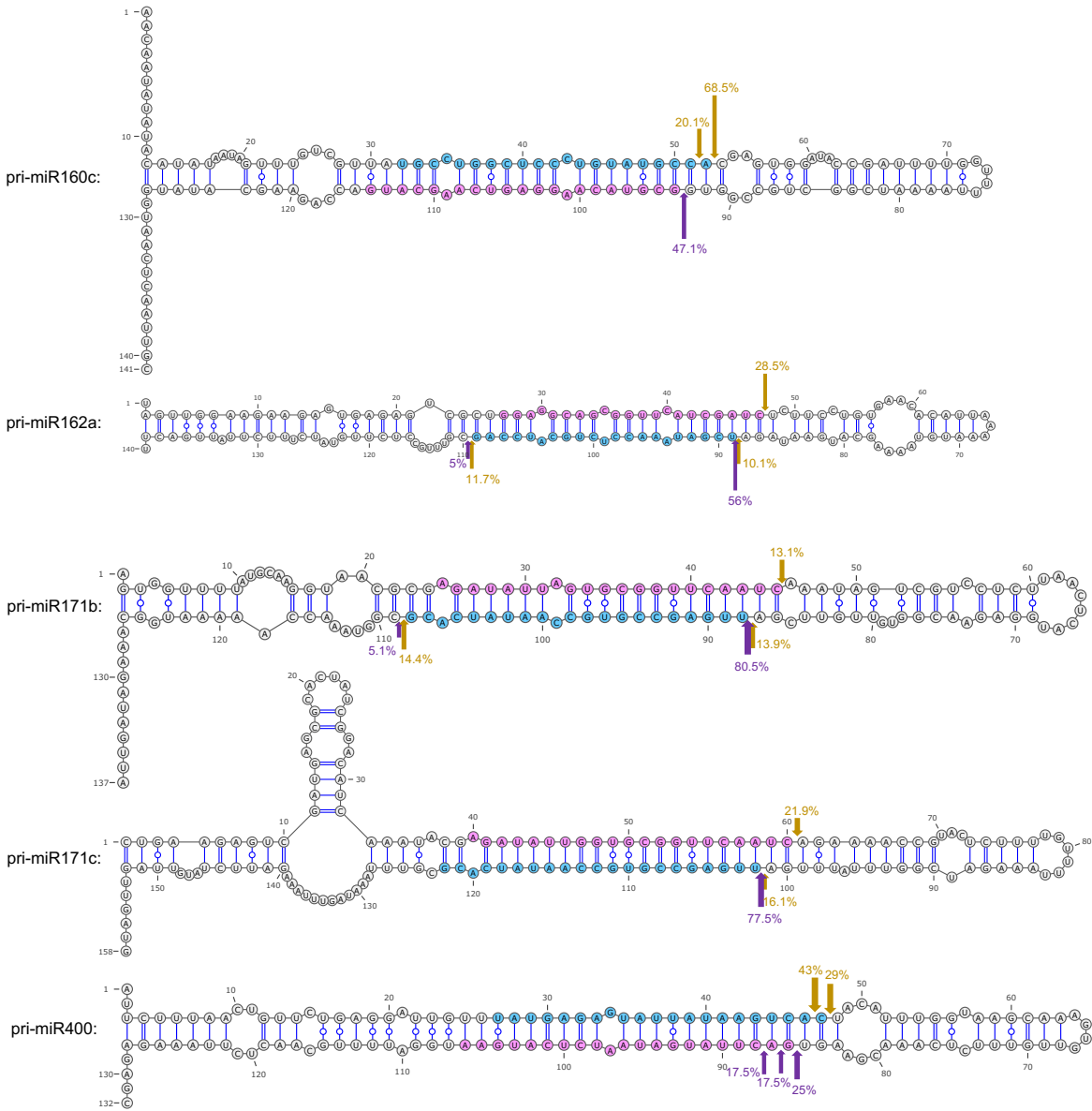
a



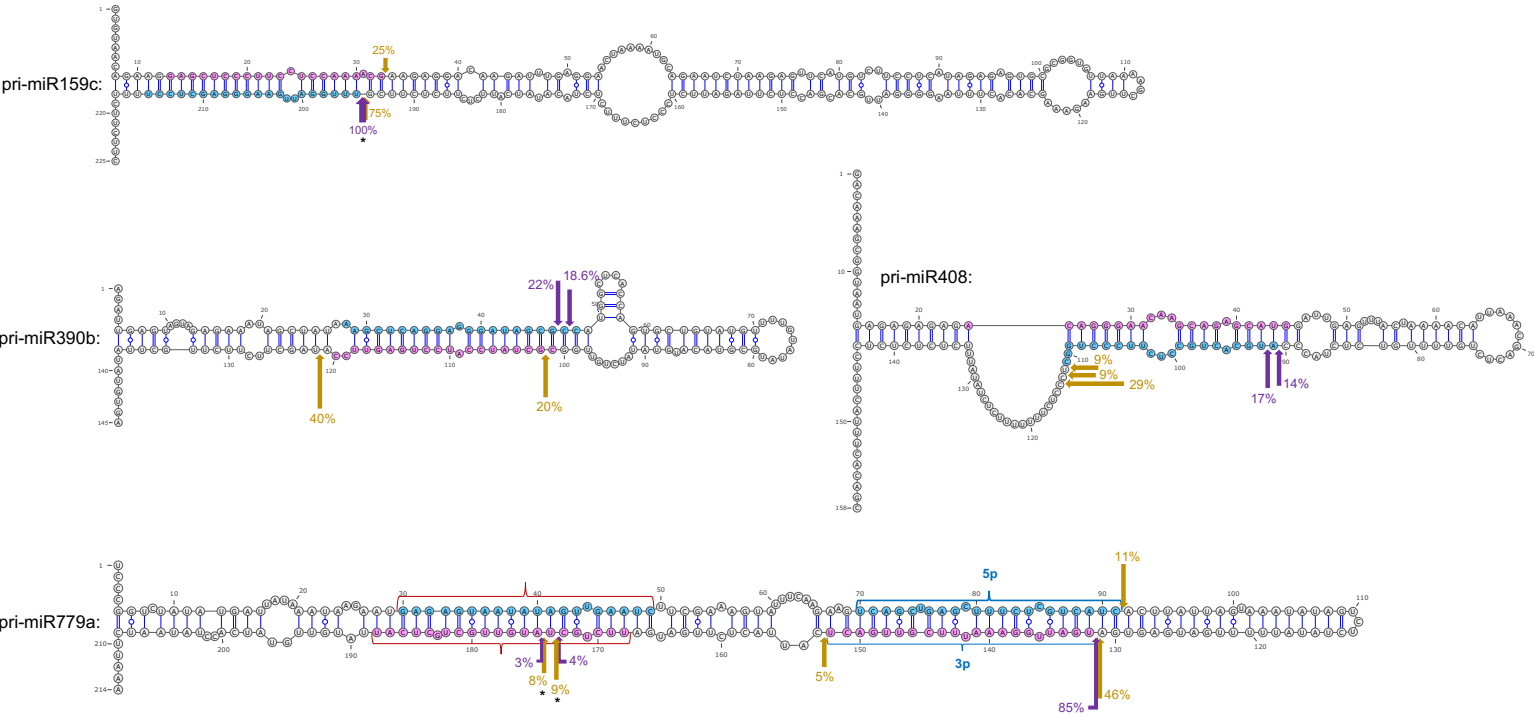
b(i) (12 re-validated)



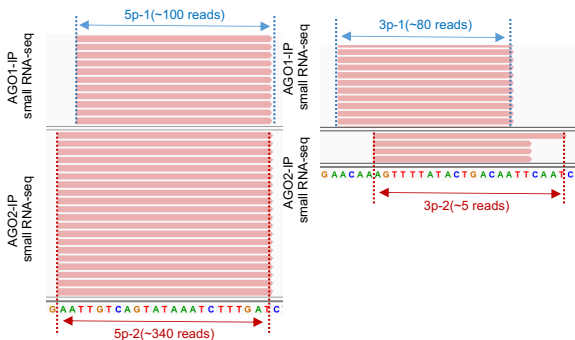
b(ii)



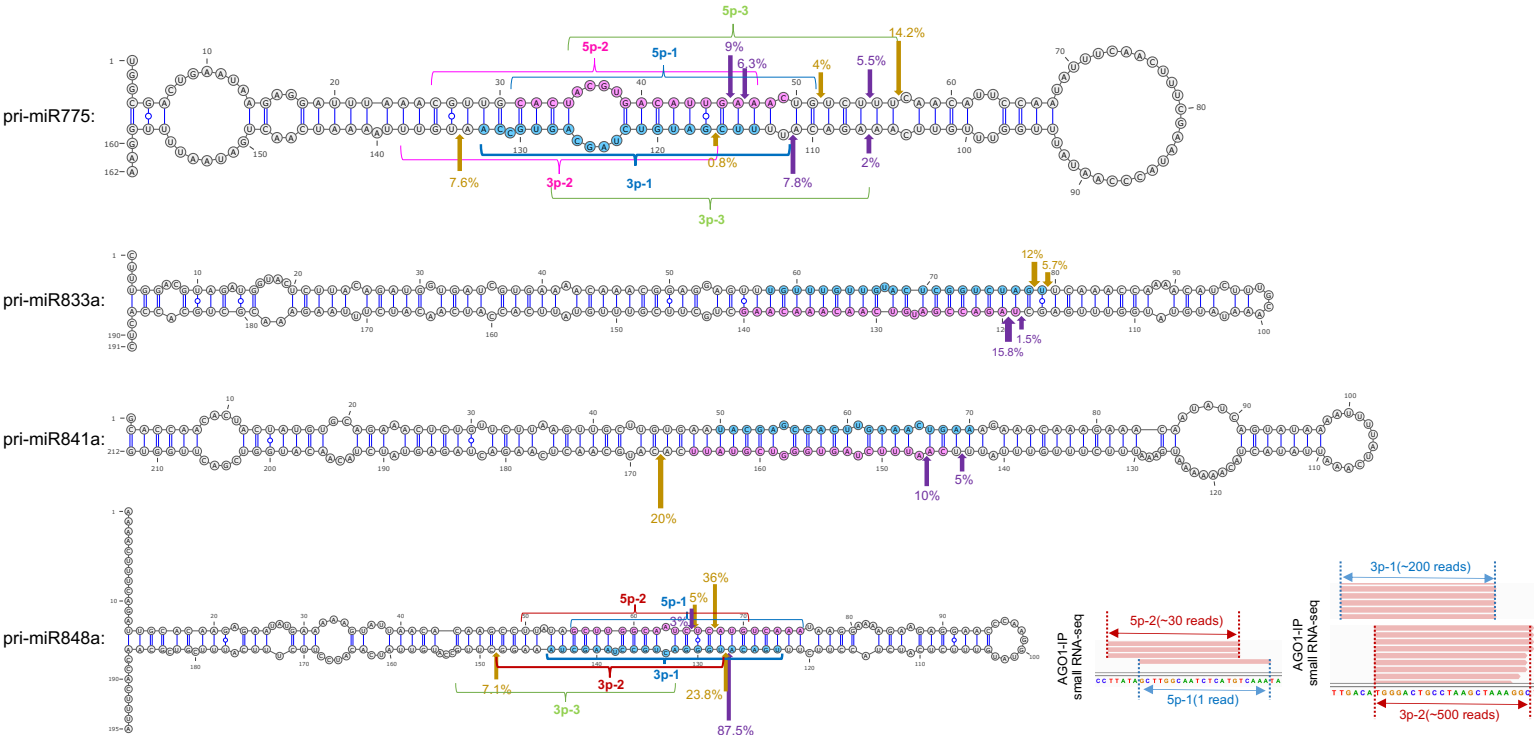
d (4 re-annotated)



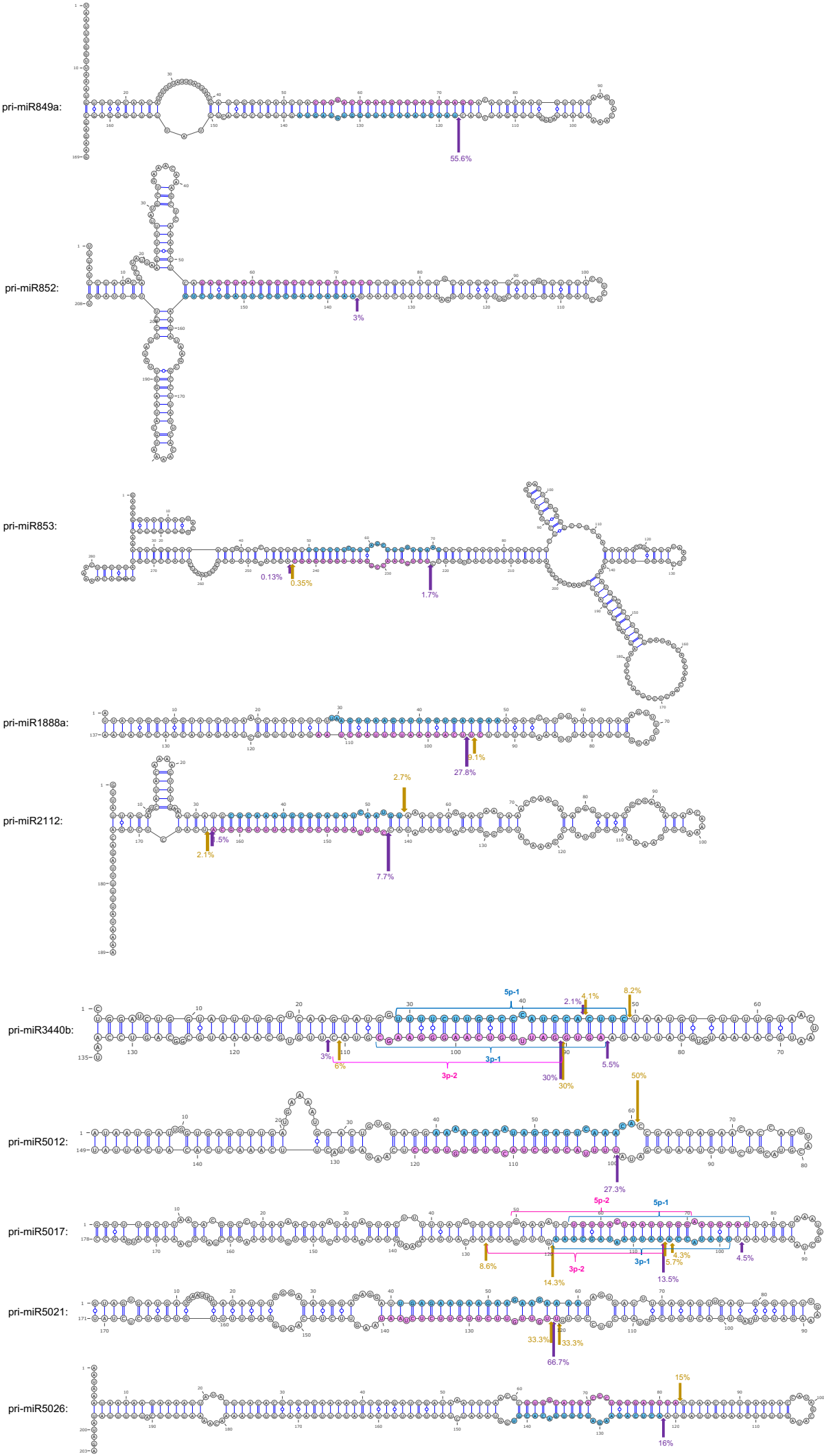
e pri-miR4245:

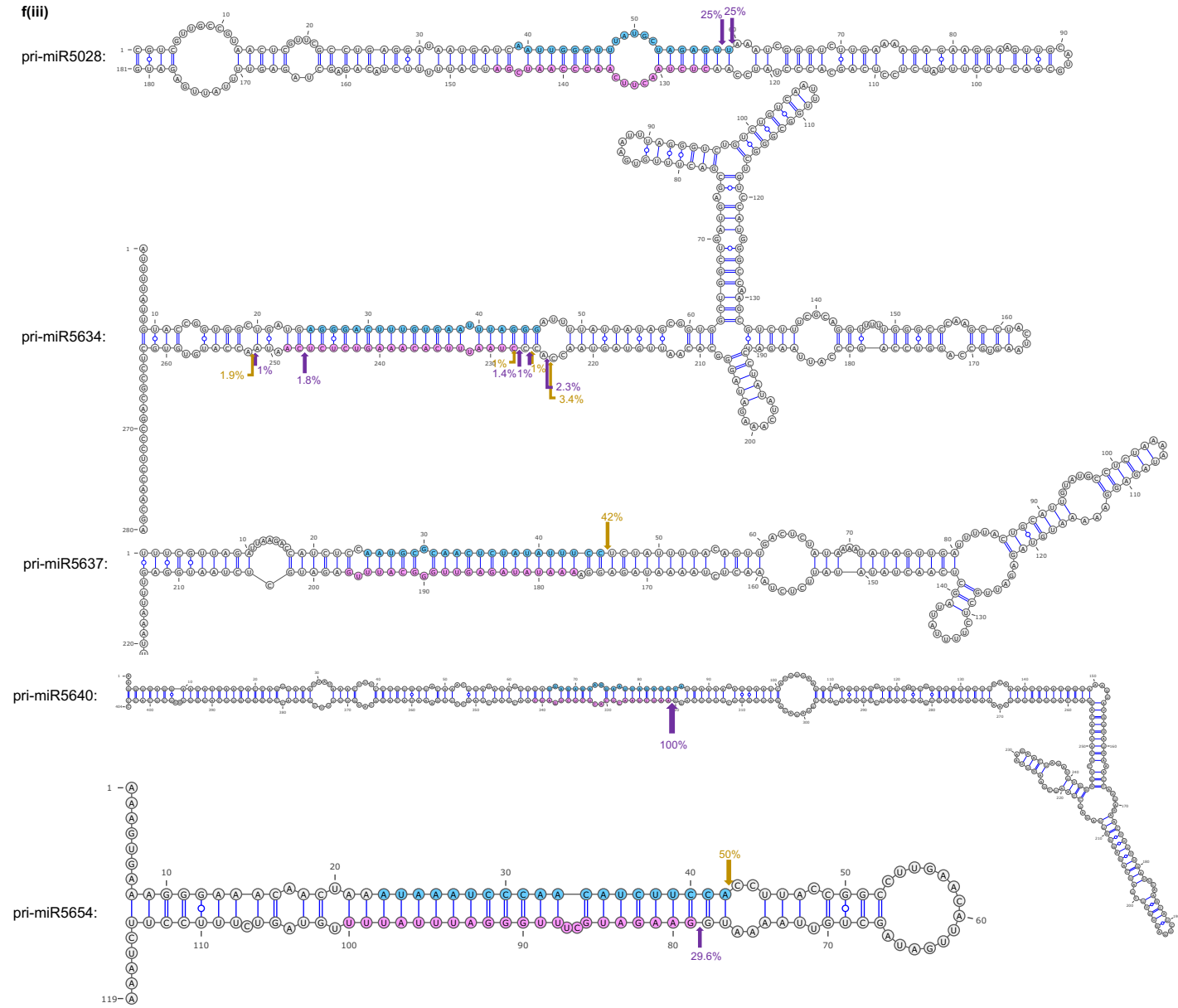


f(i) (19 newly discovered)



f(ii)





Supplementary Figure 4. Degradome sequencing reveals that 38 pri-miRNAs have LTB processing pattern.

(a) IGV graphs of reads' numbers and relative ratios at detected cleavage sites of pri-miR160a in Col-0 and semi-active DCL1 expressing lines. Related to Fig. 2f.

(b) The rest 12 LTB-processed pri-miRNAs discovered earlier are re-validated here. Be noted that there are two pages for this panel (4b(i) and 4b(ii)).

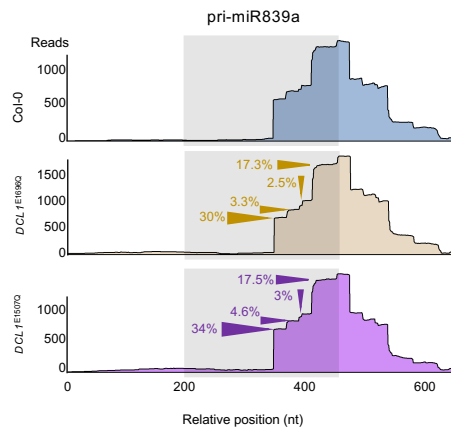
(c) IGV graphs of reads' numbers and relative ratios at detected cleavage sites of pri-miR395d in Col-0 and semi-active DCL1 expressing lines. Related to Fig. 2g.

(d) Pri-miR159c, pri-miR390b, pri-miR408 and pri-miR779a display LTB pattern rather than the earlier proposed modes. The cleavage site marked with an asterisk (*) was also discovered in WT or *fiery1*¹⁴. Of note, pri-miR408 was presented ambiguity between BTL and LTB modes, while our analysis unambiguously identified it as the LTB pattern. Pri-miR779a was reported to follow the SLTB pattern as the predicted miRNA/* is proximal to the base region and multiple times of processing from the top region would be otherwise necessary to reach the hypothetical miR779/* position (in red brackets). However, we observed the pri-miRNA also produced another set of miRNA/* (5p/3p in blue brackets) that could be recovered from AGO1-RISC with being even more abundant than the predicted miR779/*. Since the new species of miRNA/* and the first cleavage site with the highest cleavage ratios were both proximal to the top region, pri-miR779a is re-annotated to follow an LTB pattern.

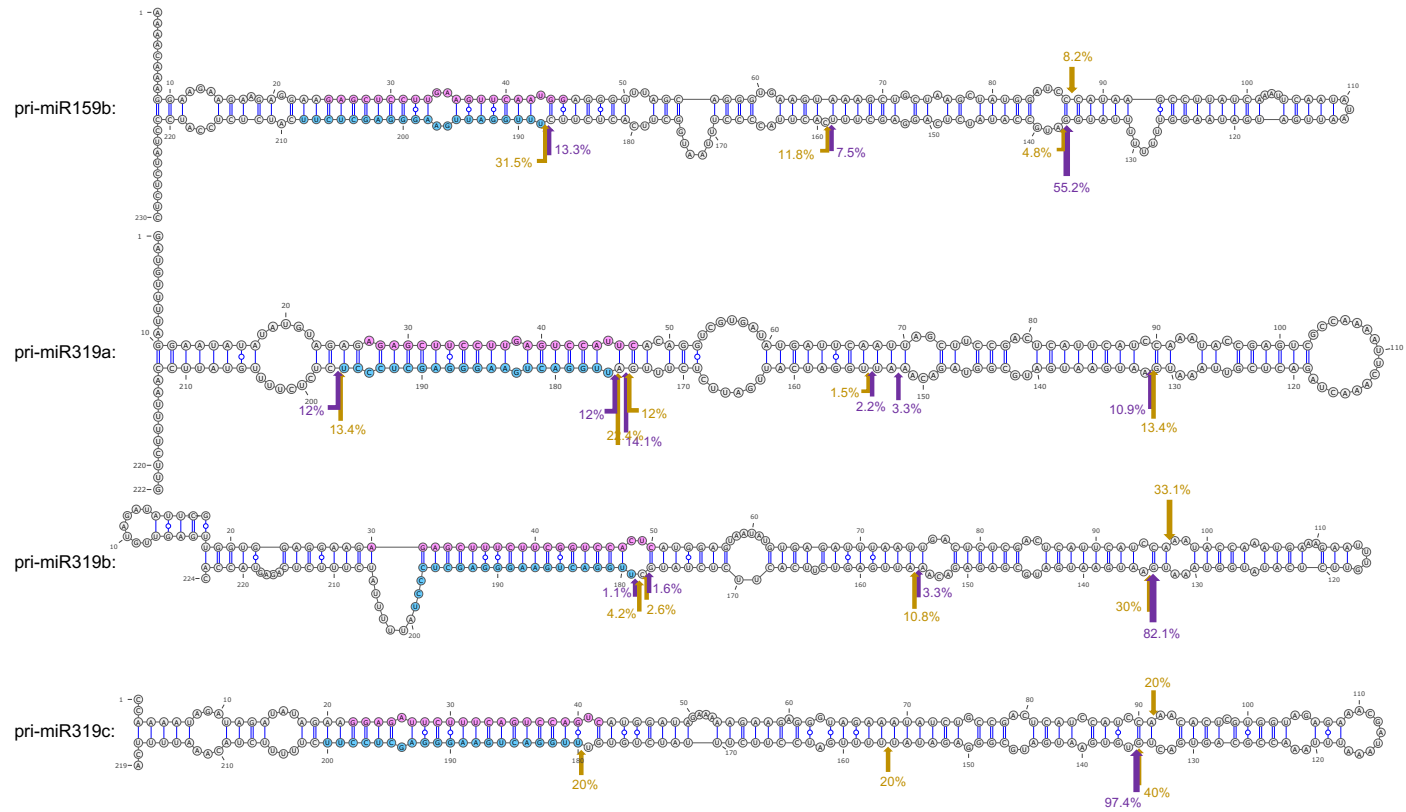
(e) IGV graphs of two species of miRNA/*s from pri-miR4245. The canonical miR4245/* (5p-1/3p-1) is sorted into AGO1 whereas the non-canonical pair (5p-2/3p-2) newly discovered here is recovered from AGO2, even with more sRNA reads. Related to Fig. 2h.

(f) 19 newly discovered LTB pri-miRNAs in our data presented in the panels in three pages (4f(i) to 4f(iii)). Notably, pri-miR775 displays three distinct LTB processing modes with comparable cleavage frequencies that produce the canonical miR775/* (5p-1/3p-1 in blue) and two new sets of miRNA/*s (5p-2/3p-2 in pink; 5p-3/3p-3 in light green). Similarly, pri-miR848a, pri-miR3440b, and pri-miR5017 all show two or three kinds of LTB processing patterns that generate new species of miRNA/*s in addition to the earlier reported miRNA/*s.

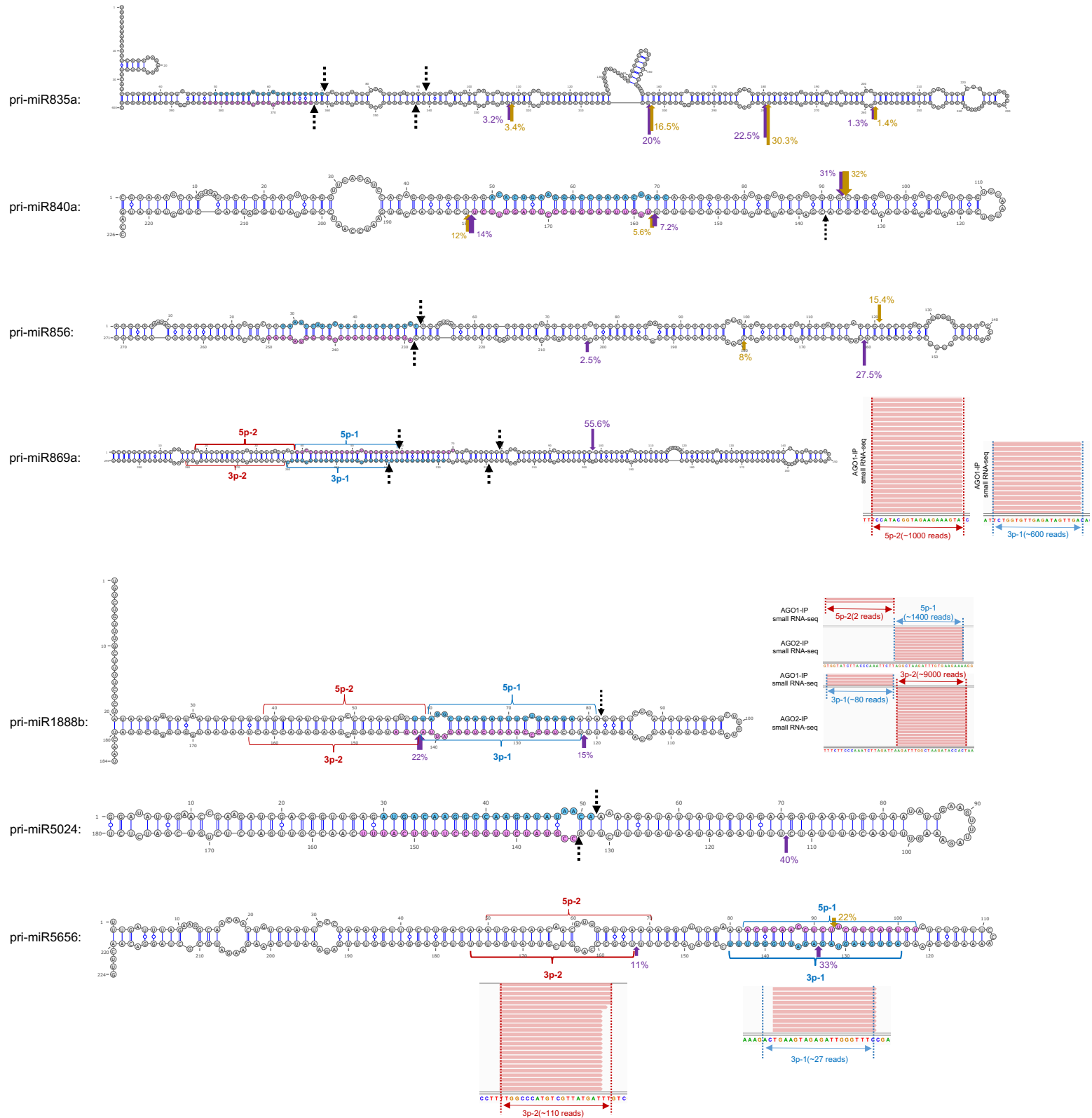
C



pri-miR159b:



d (7 newly discovered)



Supplementary Figure 5. Degradome sequencing reveals that 13 pri-miRNAs have the SLTB processing pattern.

(a) IGV graphs of reads' numbers and relative ratios at detected cleavage sites of pri-miR159a in Col-0 and semi-active DCL1 expressing lines. Related to Fig. 2i.

(b) The rest four SLTB-processed pri-miRNAs discovered earlier are re-validated here.

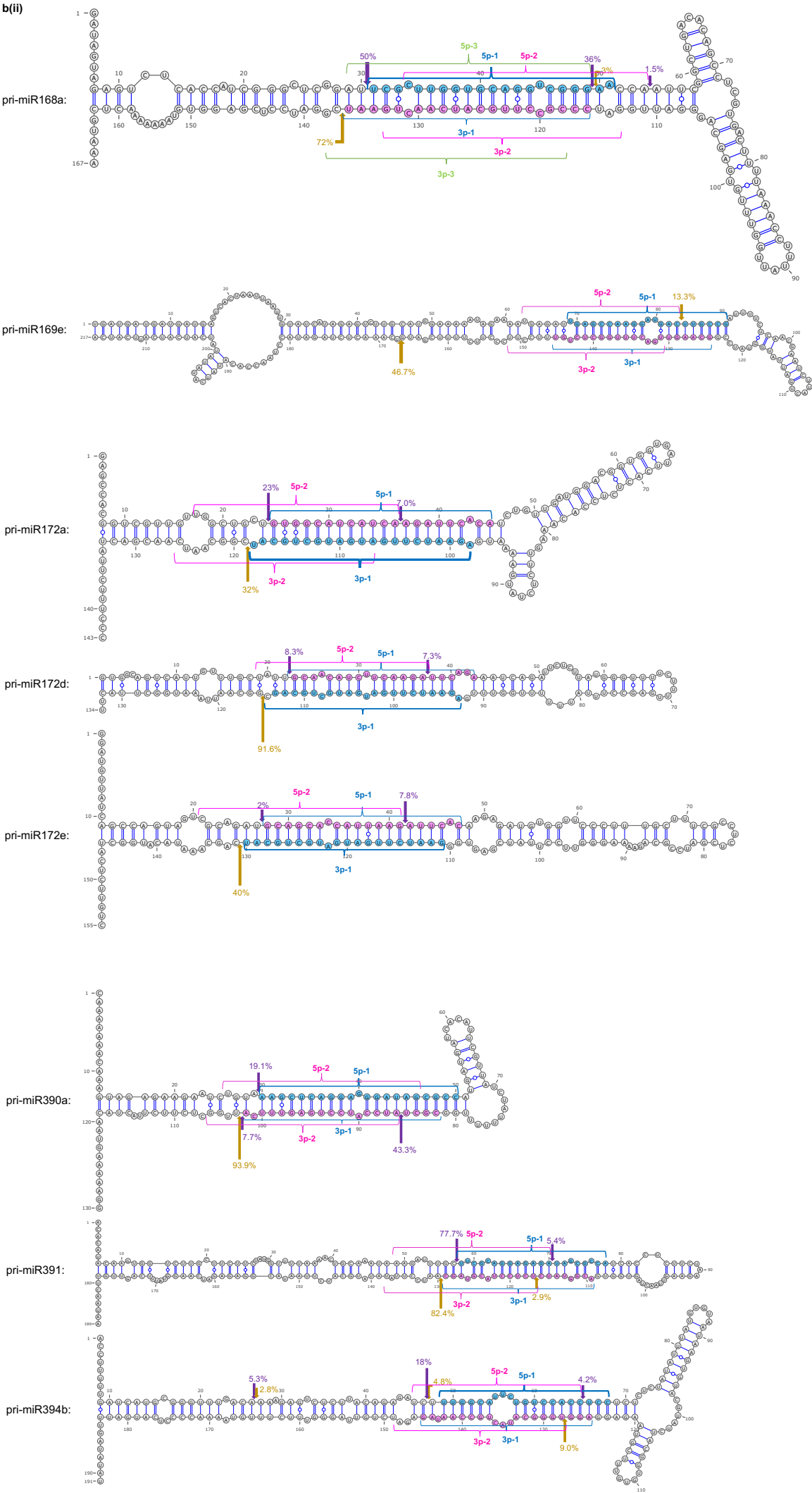
(c) IGV graphs of reads' numbers and relative ratios at detected cleavage sites of pri-miR839a in Col-0 and semi-active DCL1 expressing lines. Related to Fig. 2j.

(d) Newly identified SLTB pattern for the rest seven pri-miRNAs.

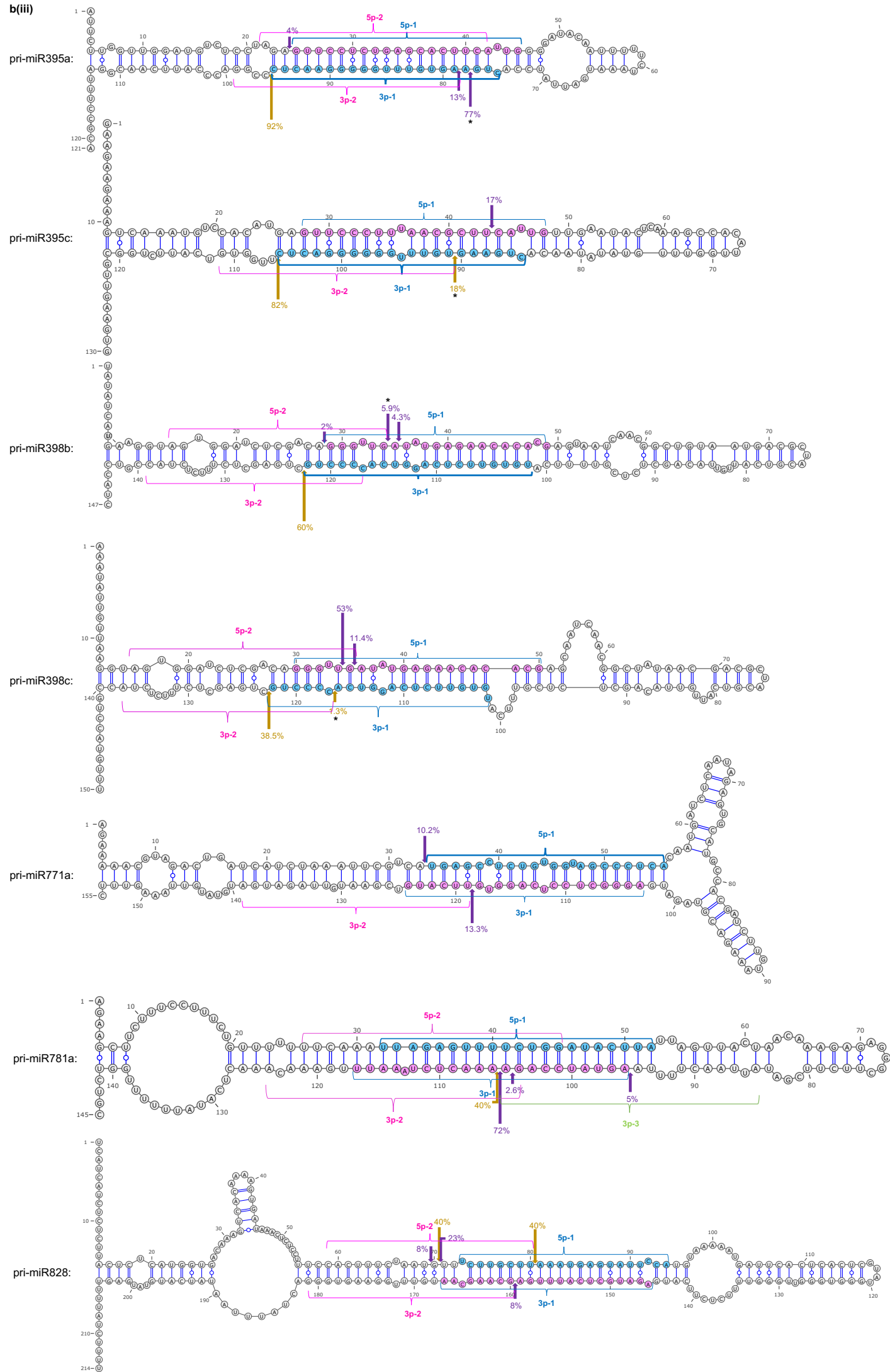
Figure 2 displays three genomic tracks for the pri-miR166a locus. The top track shows Col-0 RNA-seq reads (blue) with a peak at ~350 nt. The middle track shows DCL1-seq reads (brown) with peaks at ~250 nt (5.7%) and ~350 nt (78%). The bottom track shows DCL1-seq reads (purple) with peaks at ~250 nt (23.7%) and ~350 nt (62.1%). The x-axis is 'Relative position (nt)' from 0 to 600.

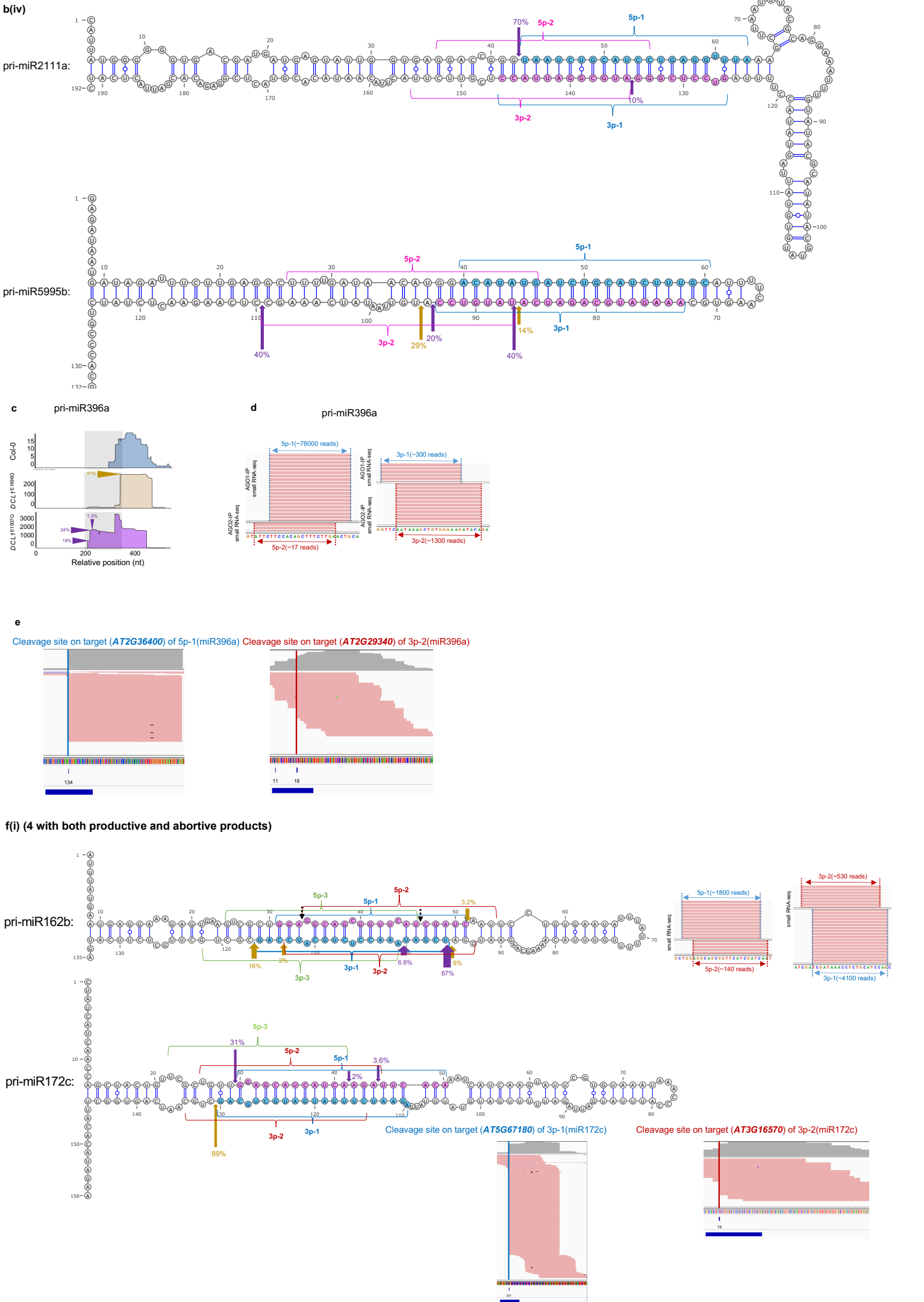
[illegible]

b(ii)

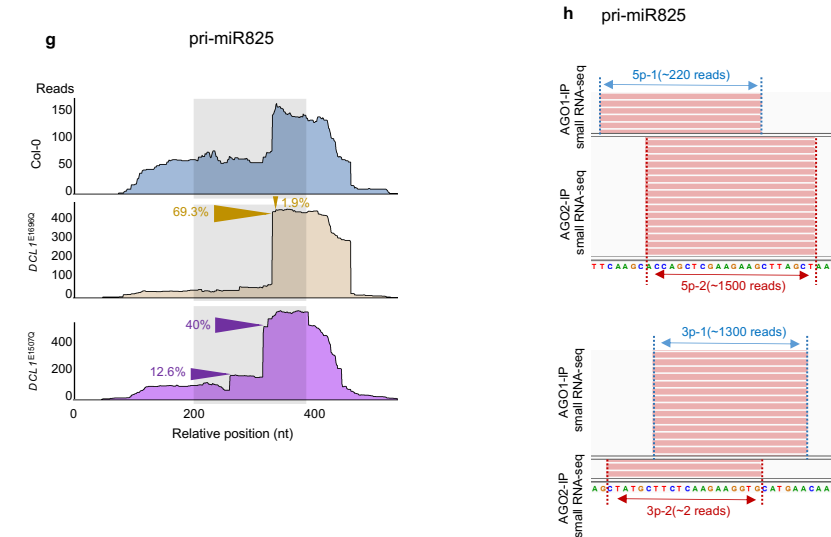
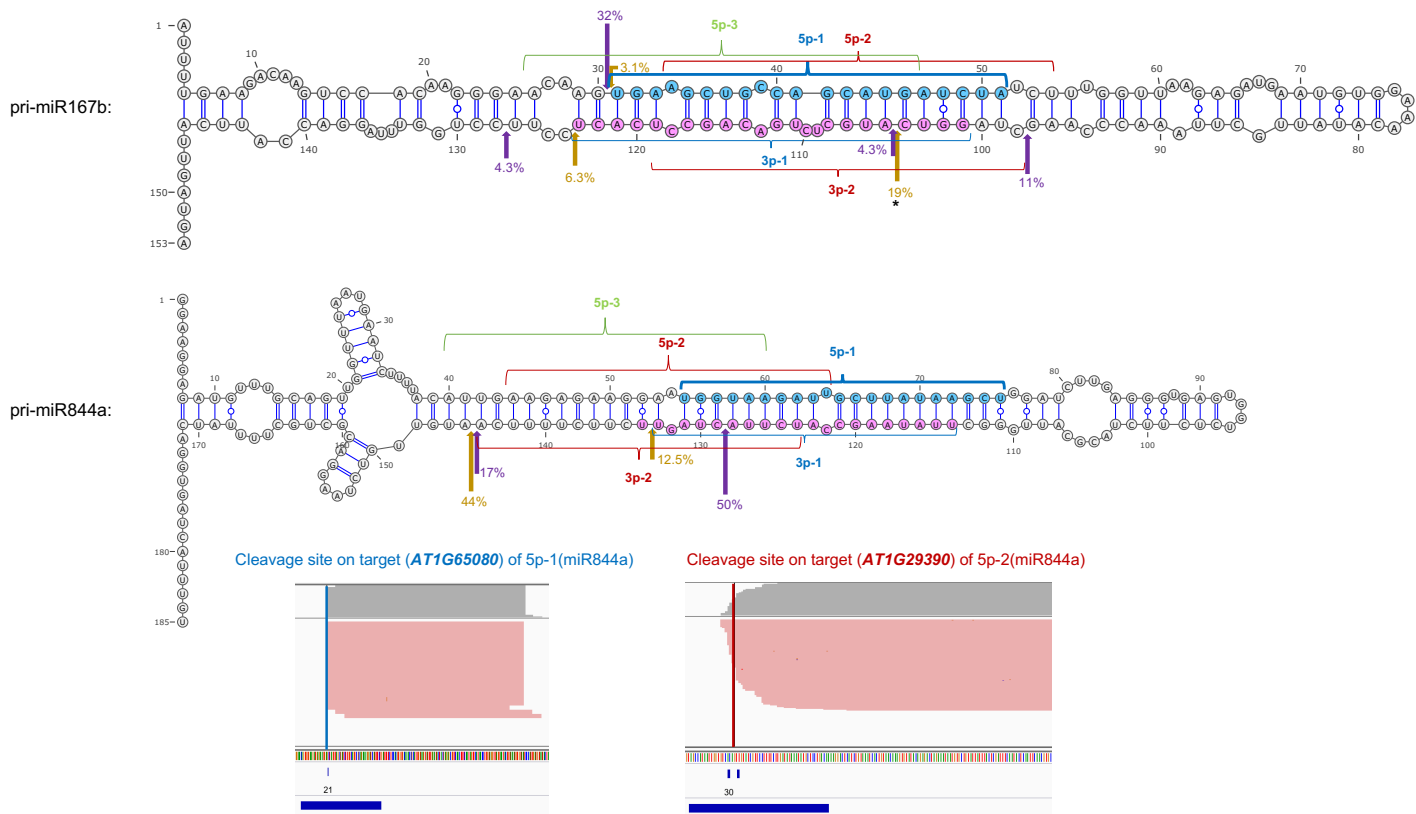


b(iii)

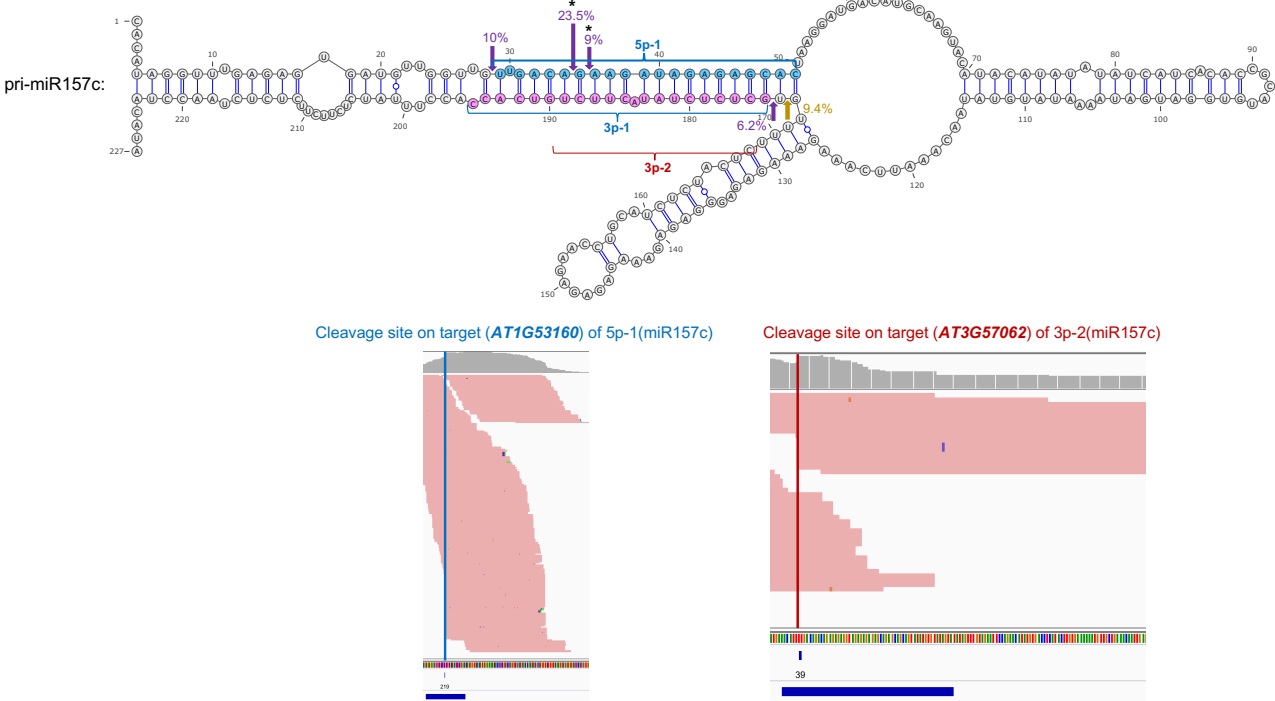




f(ii)

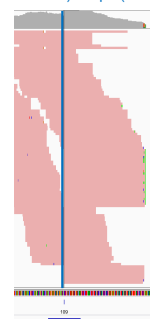
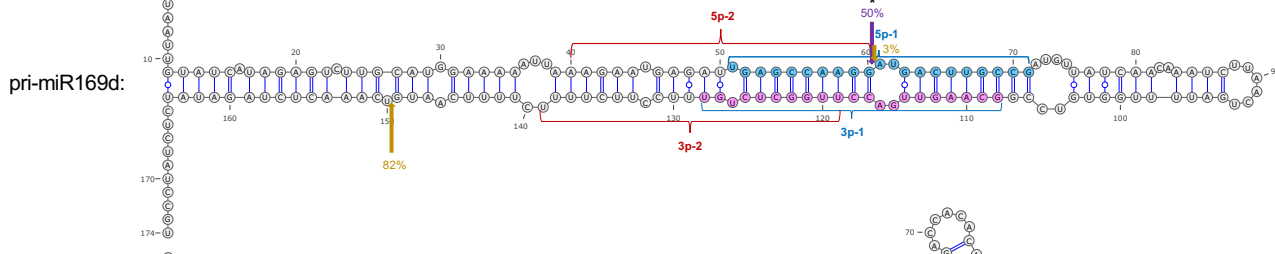


i(i) (13 with productive products)

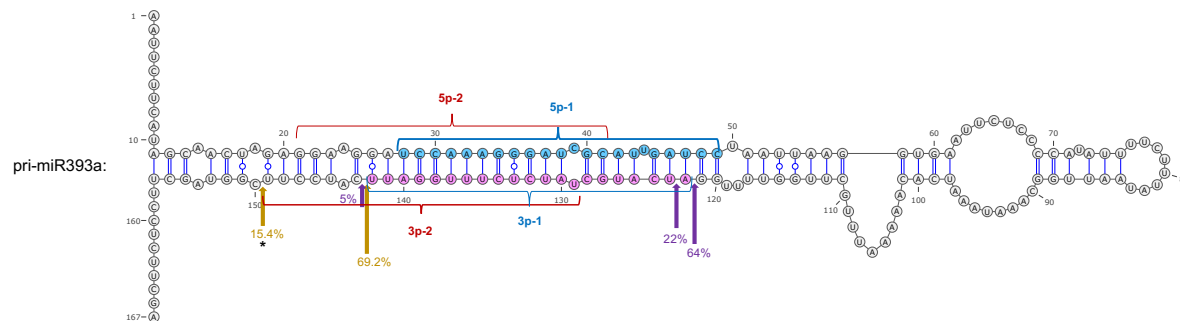
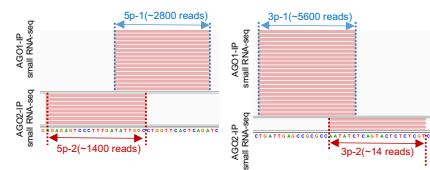
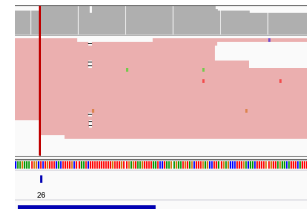
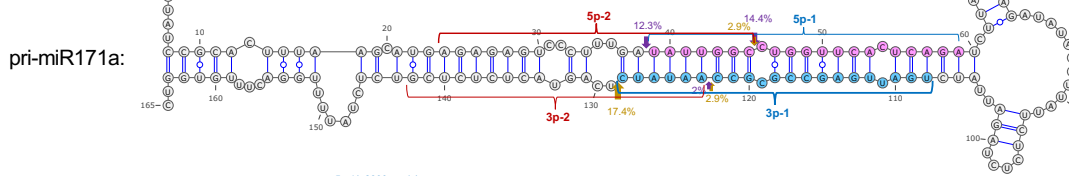


i(ii)

Cleavage site on target (*AT1G54160*) of 5p-1(miR169d)

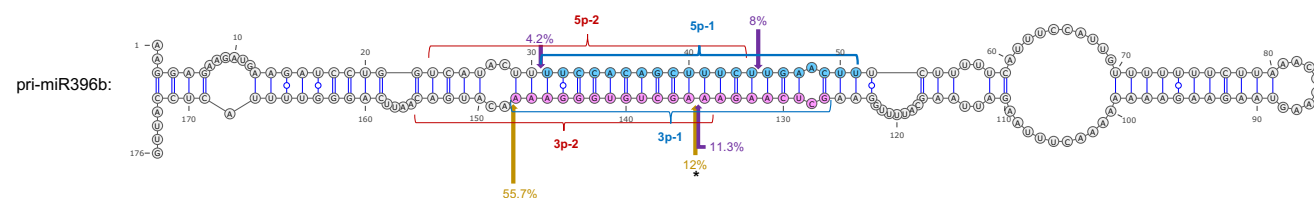
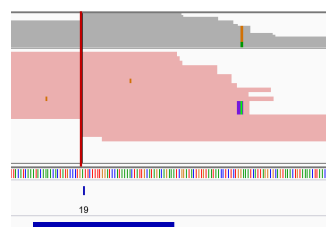
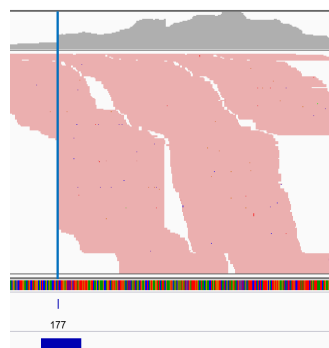


Cleavage site on target (*AT2G17660*) of 5p-2(miR169d)



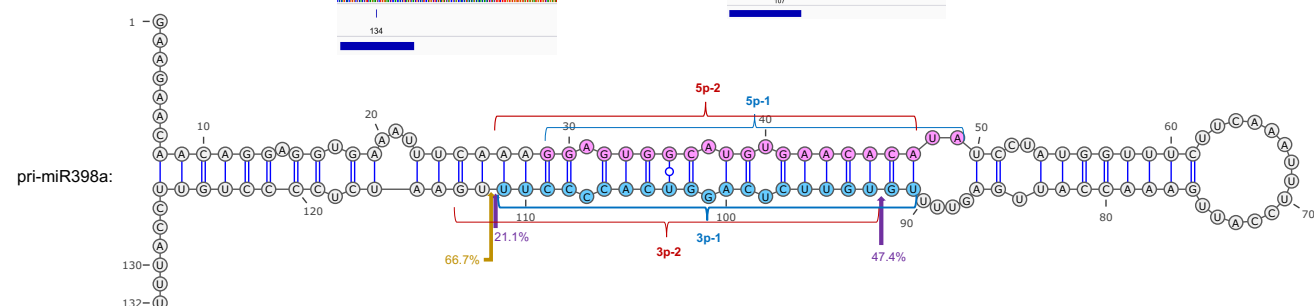
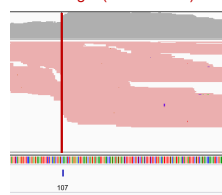
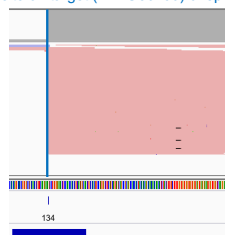
Cleavage site on target (*AT4G03190*) of 5p-1(miR393a)

Cleavage site on target (*AT1G05170*) of 5p-2(miR393a)

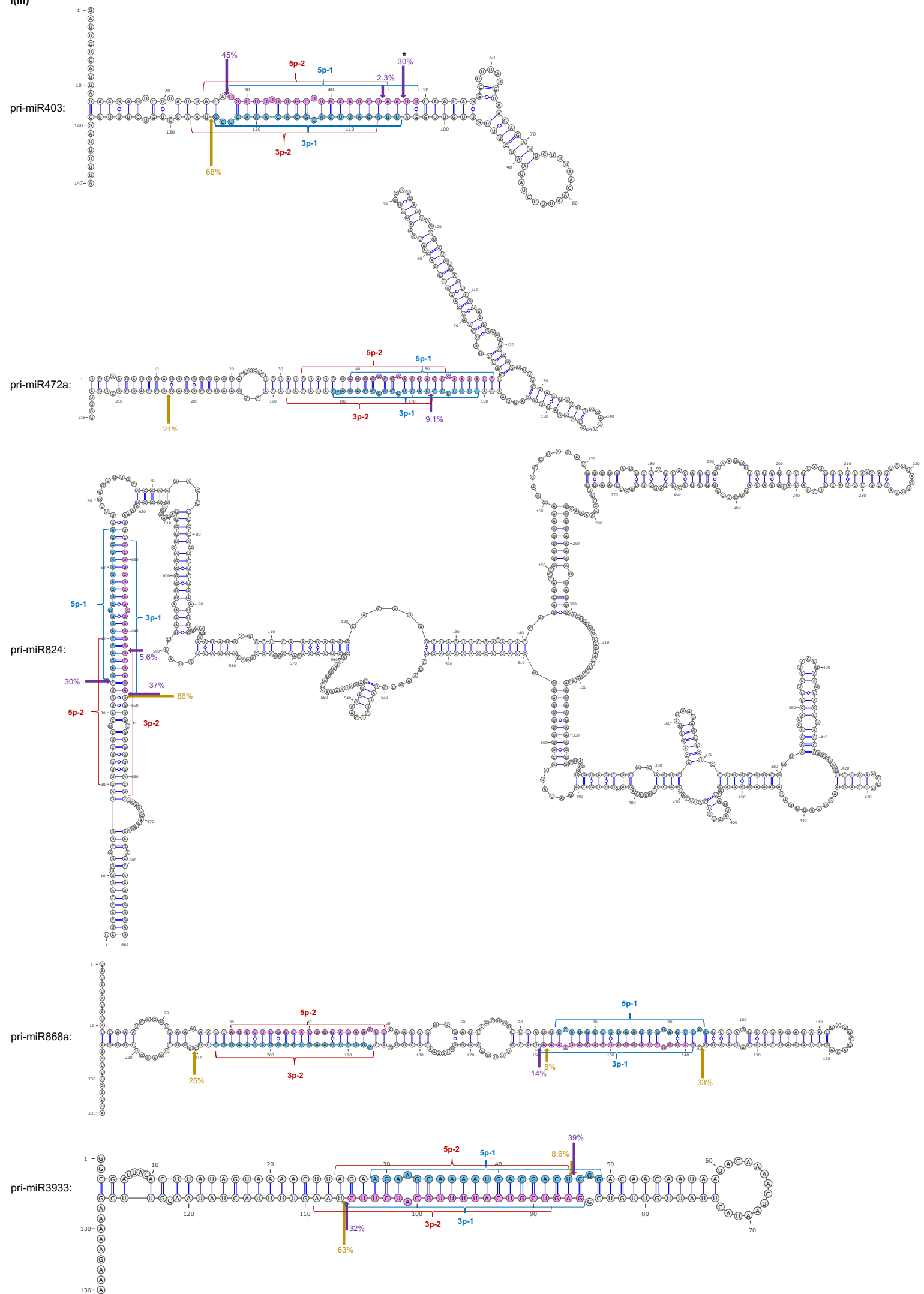


Cleavage site on target (*AT2G36400*) of 5p-1(miR396b)

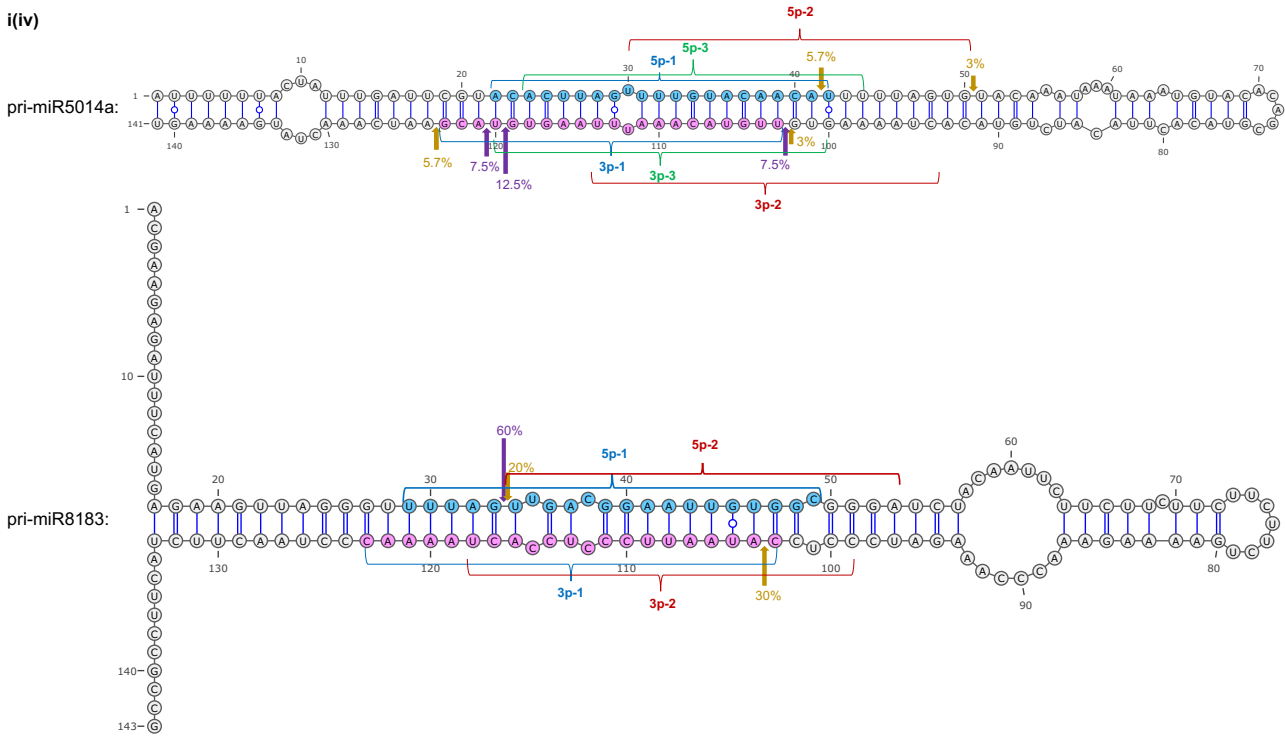
Cleavage site on target (*AT5G59700*) of 3p-2(miR396b)



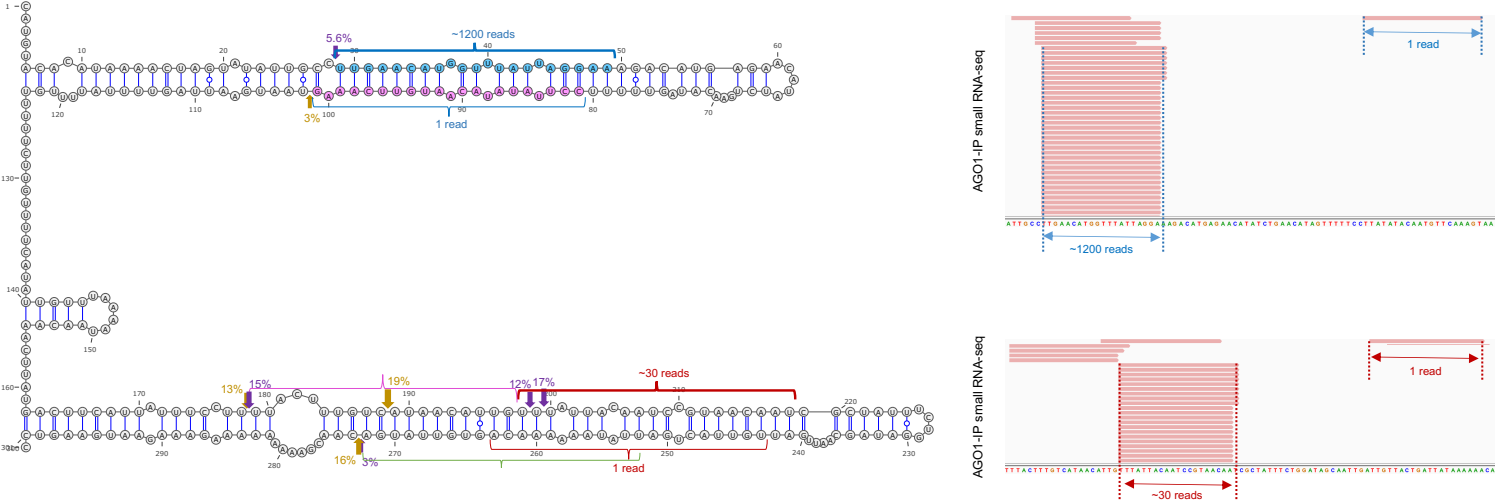
i(iii)



i(iv)



j Example of newly identified cluster of pri-miR867:



Supplementary Figure 6. Degradome sequencing reveals that 43 pri-miRNAs have bidirectional processing pattern and pri-miR867 possibly yields new miRNA clusters.

(a) IGV files of reads' numbers and relative ratios at the cleavage sites of pri-miR166a in Col-0 and semi-active DCL1 expressing lines. Related to Fig. 2k.

(b) The rest 23 pri-miRNAs show bidirectional processing patterns and can produce additional abortive products instead of the earlier reported single-directional modes that only produce one type of miRNA/*s. The cleavage site marked with an asterisk (*) was also discovered in *fiery1*¹⁴. Be noted that there are four pages for this panel (6b(i) to 6b(iv)).

(c) IGV files of reads' numbers and relative ratios at the cleavage sites of pri-miR396a in Col-0 and semi-active DCL1 expressing lines. Related to Fig. 2l.

(d) IGV graphs of two species of miRNA/*s from pri-miR396a. The canonical miR396/* (5p-1/3p-1) is sorted into AGO1 whereas the other productive pair (5p-2/3p-2) newly discovered here is recovered from AGO2. Related to Fig. 2l.

(e) Degradome-seq data verified that the newly identified 3p-2 of pri-miR396a can target a different mRNA (*AT2G29340*) for cleavage compared to annotated 5p-1. Related to Fig. 2l.

(f) The rest four pri-miRNAs display three initial processing sites and generate a new set of abortive products, and a new set of productive miRNA/* products that have their cognate mRNA targets in addition to the annotated miRNA/*s. For pri-miR162b, both the canonical 5p-1/3p-1, indicated by 67% and 3.2% cutting ratio at the upper stem, and the newly produced 5p-2/3p-2, indicated by 2% cleavage frequency at the lower stem could be recovered from sRNA-seq with adequate read counts. In addition, another set of abortive products (5p-3/3p-3 in light green) could also be produced from pri-miR162b, with a moderate frequency (6.6%). Since miR162 targets *DCL1* itself, the presence of three processing sites at the same pri-miR162b suggests that there might be sophisticated regulatory layers to fine tune the expression of this key enzyme in miRNA production in plants. The cleavage site marked with an asterisk (*) was also discovered in WT and *fiery1*¹⁴. There are two pages for this panel (6f(i) and 6f(ii)).

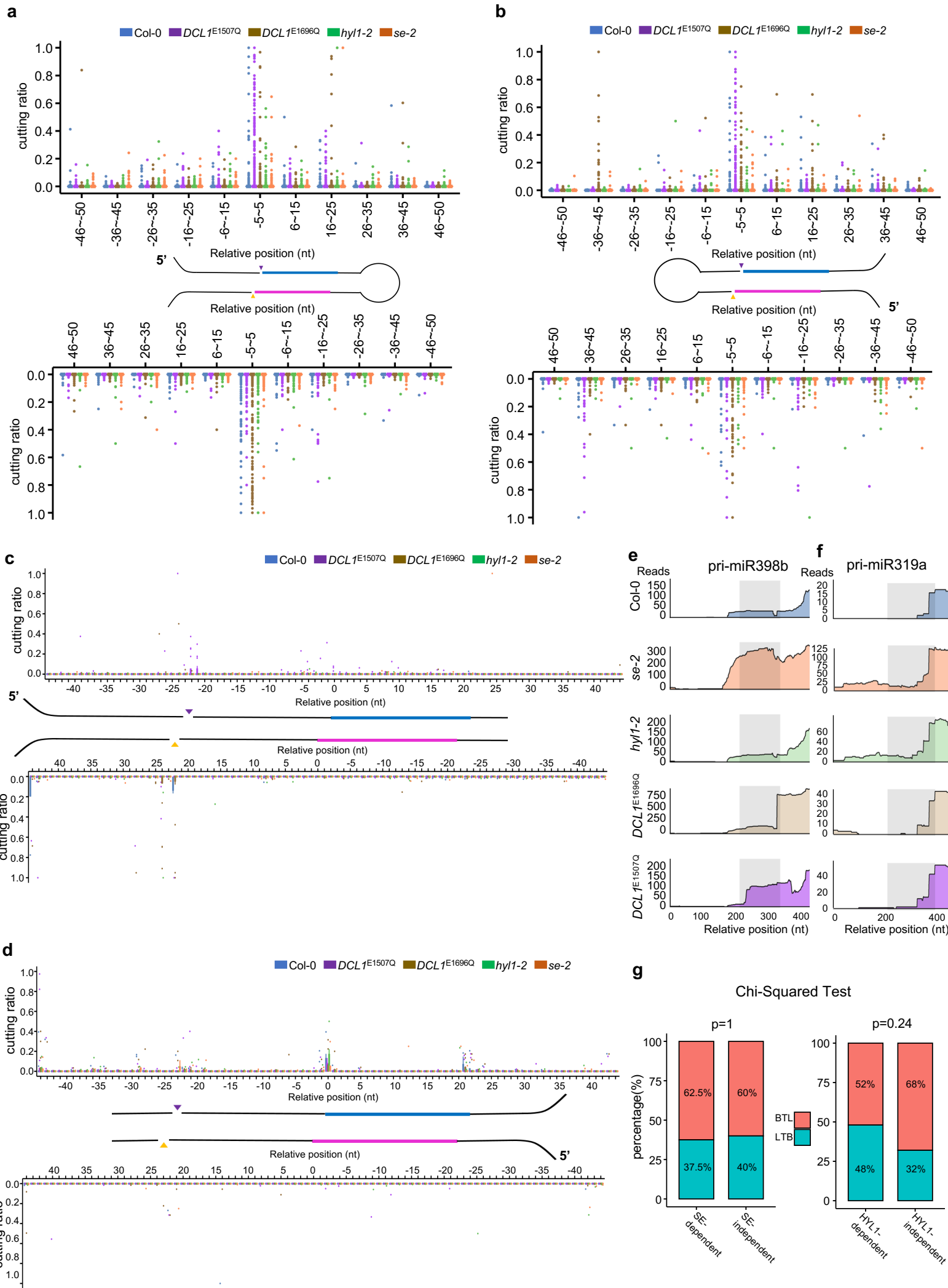
(g) IGV files of reads' numbers and relative ratios at the cleavage sites of pri-miR825 in Col-0 and semi-active DCL1 expressing lines. Related to Fig. 2m.

(h) IGV graphs of two species of miRNA/*s from pri-miR825. The canonical miR825/* (5p-1/3p-1) is sorted into AGO1 whereas the other productive pair (5p-2/3p-2) newly discovered here is recovered from AGO2, with more sRNA reads than canonical one. Related to Fig. 2m.

(i) Newly identified bidirectional processing patterns for the rest 13 pri-miRNAs and resultant productive products with some new targets included. The cleavage site marked with an asterisk (*) was also discovered in *fiery1*¹⁴. Be noted that there are four pages for this panel (6i(i) to 6i(iv)).

(j) Pri-miR867 is one of the newly identified cases that undergo BTL processing and/or possibly additional processing of the adjacent arm to produce new but low abundant miRNA clusters. IGV graphs of sRNAs reads recovered from AGO1-IP are shown (right panel). Be noted: pri-miR867 was detected with several cutting sites with high cutting ratios at downstream of the transcript. RNAfold predicted that pri-miR867 had two bilateral fold-back hairpin structures: one fold-back structured segment appeared to be cut from base to loop, producing miR867/* that can be recovered from AGO1-RISC. The other fold-back segment, however, displayed a SBTL pattern with cutting ratios of 15% and 13% that generates intermediate product (enclosed by the pink bracket). Meanwhile, this segment also underwent a BTL mode with cutting ratios 12% and 17%, resulting in new products (enclosed by the red brackets) that can be found in AGO1-RISC, despite of the relatively low read counts. Furthermore, an additional BTL processing with cutting frequencies of 16% and 19% was also detected but yielded an abortive product (enclosed by the light green bracket). This scenario bears resemblance to the clustered miRNAs observed in the animal system, suggesting the presence of additional but non-canonical DCL1 substrates in plants.

Supplementary Fig. 7



Supplementary Figure 7. SE and HYL1 show different impacts on pri-miRNA processing.

(a, b) Boxplots show the cutting ratios for BTL **(a)** and LTB **(b)** pri-miRNAs in Col-0 (blue bar), *DCL1*^{E1507Q} (purple bar), *DCL1*^{E1696Q} (yellow-brown bar), *hyl1-2* (green bar) and *se-2* (orange bar). In the top panel, position “0” is defined as the first cutting sites in *DCL1*^{E1507Q}, same sites as the purple arrowhead labeled in the pri-miRNA cartoon. In the bottom panel, position “0” is defined as the first cutting sites in *DCL1*^{E1696Q}, the same sites as the yellow arrowhead labeled in the pri-miRNA cartoon. The blue and pink regions in the pri-miRNA cartoon represent miRNA/* duplex. Data are from two biological replicates for each sample. Centres of the boxes represent the median values. Upper bound and lower bound show the first and the third quartiles respectively. Whiskers indicate data within 1.5× the interquartile range of both quartiles. Data points at the ends of whiskers represent outliers.

(c, d) Boxplots show the cutting ratios for SBTL **(c)** and SLTB **(d)** pri-miRNAs around base/top and duplex regions in Col-0 (blue bar), *DCL1*^{E1507Q} (purple bar), *DCL1*^{E1696Q} (yellow-brown bar), *hyl1-2* (green bar) and *se-2* (orange bar). In both top and bottom panels, position “0” is defined as the first nucleotides of duplex region, the purple and yellow arrowheads labeled in the pri-miRNA cartoon represent first cleavage sites. The blue and pink regions represent miRNA/* duplex. Data are from two biological replicates for each sample. Centres of the boxes represent the median values. Upper bound and lower bound show the first and the third quartiles respectively. Whiskers indicate data within 1.5× the interquartile range of both quartiles. Data points at the ends of whiskers represent outliers.

(e, f) Schematic diagram of reads accumulation to present two categories of pri-miRNAs for which, the processing is dependent **(e)** or independent **(f)** on SE and HYL1.

(g) Chi-Squared Test for SE/HYL1 dependency on pri-miRNA processing patterns. The *P* value from 88 SE-dependent (55 for BTL and 33 for LTB) and 20 SE-independent (12 for BTL and 8 for LTB) pri-miRNAs is 1. The *P* value from 77 HYL1-dependent (40 for BTL and 37 for LTB) and 25 HYL1-independent (17 for BTL and 8 for LTB) pri-miRNAs is 0.24. The data are from two biological replicates for *se-2* and *hyl1-2*.

Unprocessed western blots for Supplementary Fig. 1a.

