Current Biology, Volume *22*

Supplemental Information

The *Arabidopsis* Nucleotidyl Transferase HESO1 Uridylates Unmethylated Small RNAs to Trigger Their Degradation

Yuanyuan Zhao, Yu Yu, Jixian Zhai, Vanitharani Ramachandran, Thanh Theresa Dinh, Blake C. Meyers, Beixin Mo, and Xuemei Chen

Supplemental Inventory

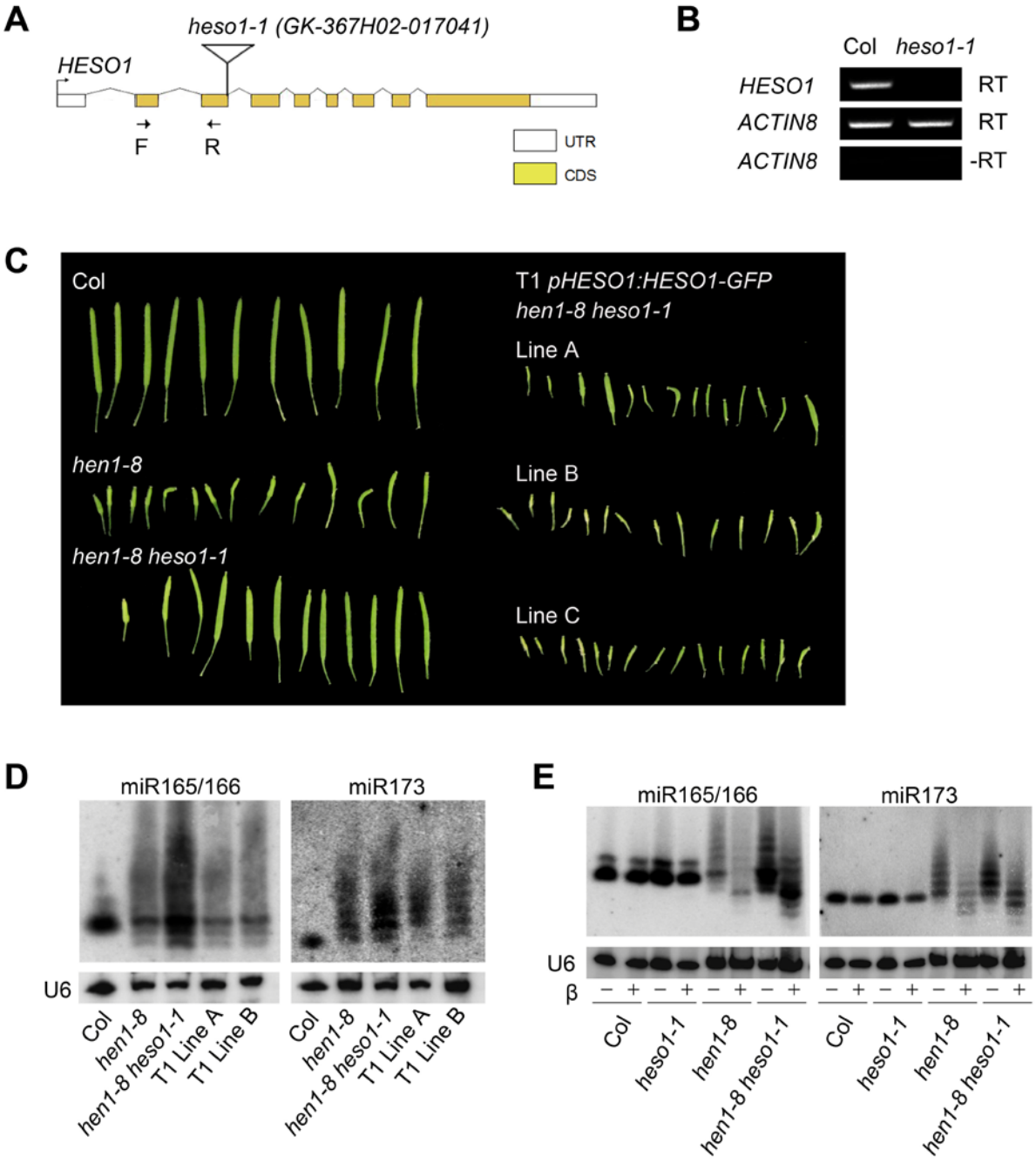

A
heso1-1 (GK-367H02-017041)
HESO1
F
R
UTR
CDS
B
Col
heso1-1
HESO1
ACTIN8
ACTIN8
RT
RT
-RT
C
Col
hen1-8
hen1-8 heso1-1
T1 pHESO1:HESO1-GFP
hen1-8 heso1-1
Line A
Line B
Line C
D
miR165/166
miR173
U6
Col
hen1-8
hen1-8 heso1-1
T1 Line A
T1 Line B
E
miR165/166
miR173
U6
β
− + − + − + − +
Col
heso1-1
hen1-8
hen1-8 heso1-1

Figure S1. Characterization of a Loss-of-Function Mutant Allele of *HESO1*, Related to Figure 1

(A) The *HESO1* gene structure and the T-DNA insertion site in the *heso1-1* mutant. Rectangles represent exons and lines represent introns. The T-DNA insertion is indicated by a triangle. The catalog number for the T-DNA mutant at the Arabidopsis Stock Center is in the parentheses. The arrow on the far left indicates the transcription start site. The arrows in the coding region mark the positions of primers used for RT-PCR in (B).

(B) RT-PCR to examine the *HESO1* transcript in wild type (Col) and *heso1-1*. *ACTIN8* was used as a loading control. Although the primers in *HESO1* were 5' to the T-DNA insertion site, no transcripts were detected in the *heso1-1* mutant, suggesting that *heso1-1* is likely a null allele.

(C) Siliques from wild-type (Col), *hen1-8*, and *hen1-8 heso1-1* plants as well as three independent *hen1-8 heso1-1* T1 transgenic lines harboring the *HESO1* genomic DNA. The *heso1-1* mutation largely restored the fertility of *hen1-8* plants, and the enhanced fertility in the *hen1-8 heso1-1* double mutant was abrogated by the *HESO1* transgene.

(D) Northern blots of miR165/166 and miR173 in wild type (Col), *hen1-8*, *hen1-8 heso1-1* and two independent T1 lines (T1 line A and line B) of *hen1-8 heso1-1* harboring the *pHESO1:HESO1-GFP* transgene. The abundance of the two miRNAs was decreased in the two T1 lines compared to *hen1-8 heso1-1*, indicating that *pHESO1:HESO1-GFP* rescued the molecular phenotypes of *heso1-1*.

(E) The *heso1-1* mutation does not rescue the miRNA methylation defects of *hen1-8*. Total RNAs were subjected (+) or not (-) to the β-elimination reactions followed by northern blotting. In *hen1-8*, the lack of 2'-*O*-methylation of the miRNAs rendered the miRNAs susceptible to the reactions such that a mobility shift was observed between the treated and untreated samples. In wild type (Col) and *heso1-1*, 2'-*O*-methylation protected the miRNAs from the chemical reactions and no mobility shifts were observed. In *hen1-8 heso1-1*, mobility shifts were observed, indicating that the miRNAs lacked 2'-*O*-methylation. Note that 50 μg of total RNA was used for *hen1-8* and *hen1-8 heso1-1* whereas 5 μg was used for wild type (Col) and *heso1-1*. The general reduction in miRNA levels after β-elimination was due to loss incurred during the procedure.

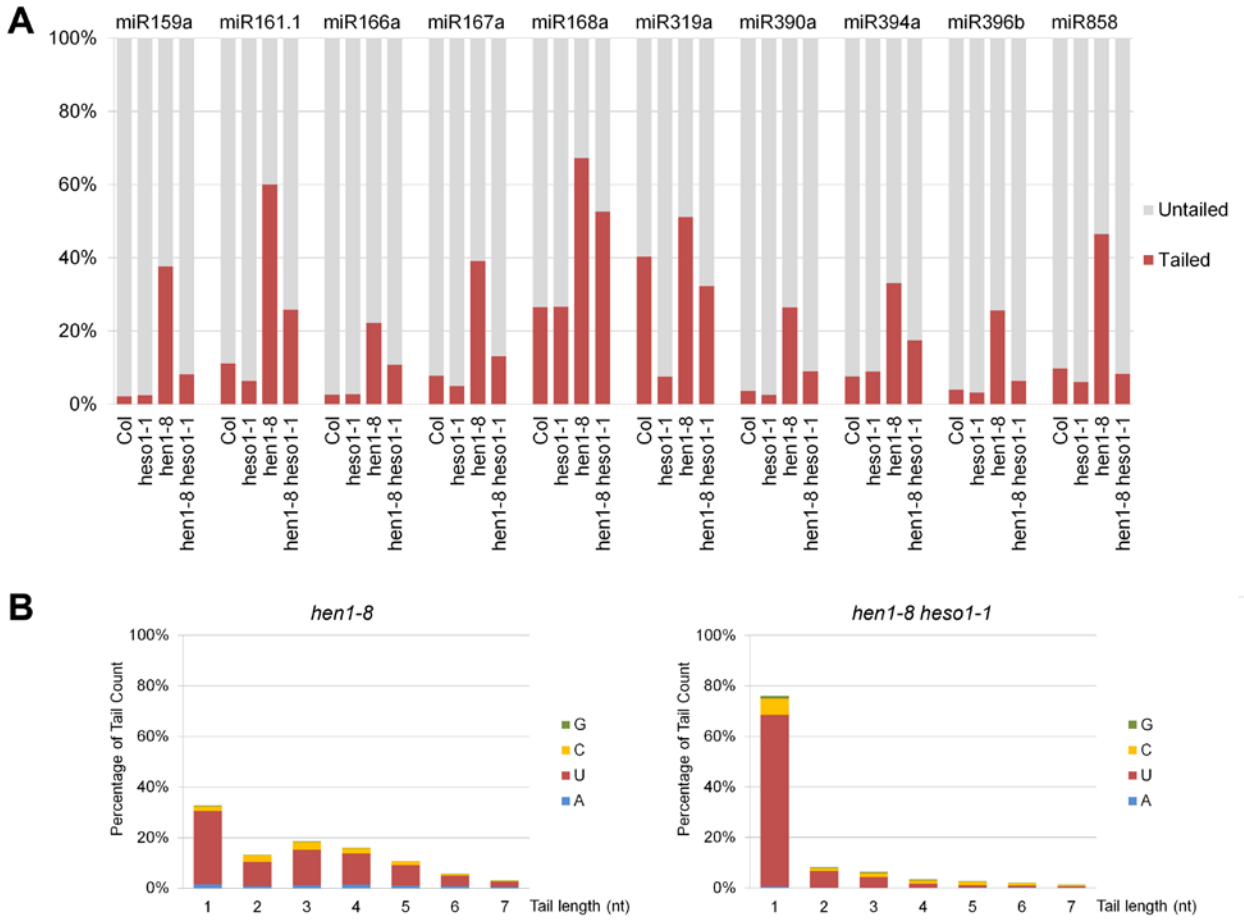


Figure S2. Analysis of 3' Tailing of miRNAs in *hen1-8* and *hen1-8 heso1-1* Plants, Related to Figure 2

High throughput sequencing was conducted for wild-type, *heso1-1*, *hen1-8*, and *hen1-8 heso1-1* genotypes with two biological replicates for each genotype. Small RNA reads corresponding to known miRNAs were categorized into four classes: full-length (class 1), tailed only (full-length reads plus tails) (class 2), truncated only (class 3), and truncated and tailed (class 4).

(A) The proportions of tailed and untailed small RNA reads for each of the ten selected miRNAs in *hen1-8* and *hen1-8 heso1-1*. The proportions of tailed and untailed species were calculated as %(sum of classes 2 and 4 read numbers/total read number) and %(sum of classes 1 and 3 read numbers/total read number), respectively. Results shown were calculated from the first biological replicate.

(B) Tail length distribution and nucleotide frequencies in the tails of selected miRNAs in *hen1-8* and *hen1-8 heso1-1*. Tailing profiles for four miRNAs from the second biological replicate are shown. For each miRNA, all tailed reads (classes 2 and 4) were analyzed for frequencies of tail length. The proportions of miRNA species with tails of 1-7 nt are shown. Nucleotide frequencies in the tails were calculated as %(number of a nucleotide in the tail/tail length).

Figure S3-1

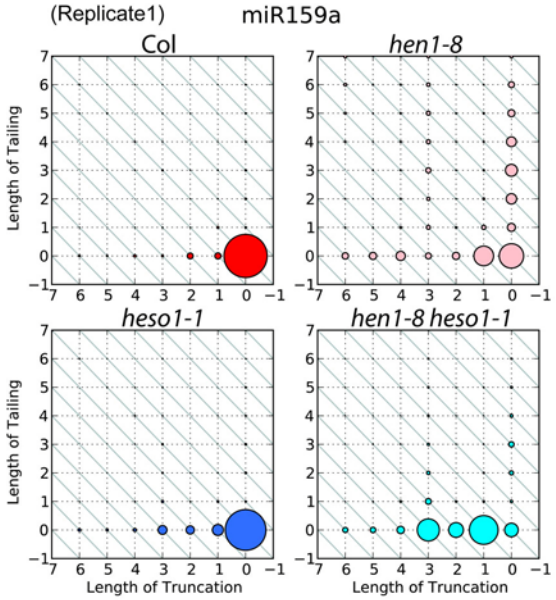


Figure S3-2

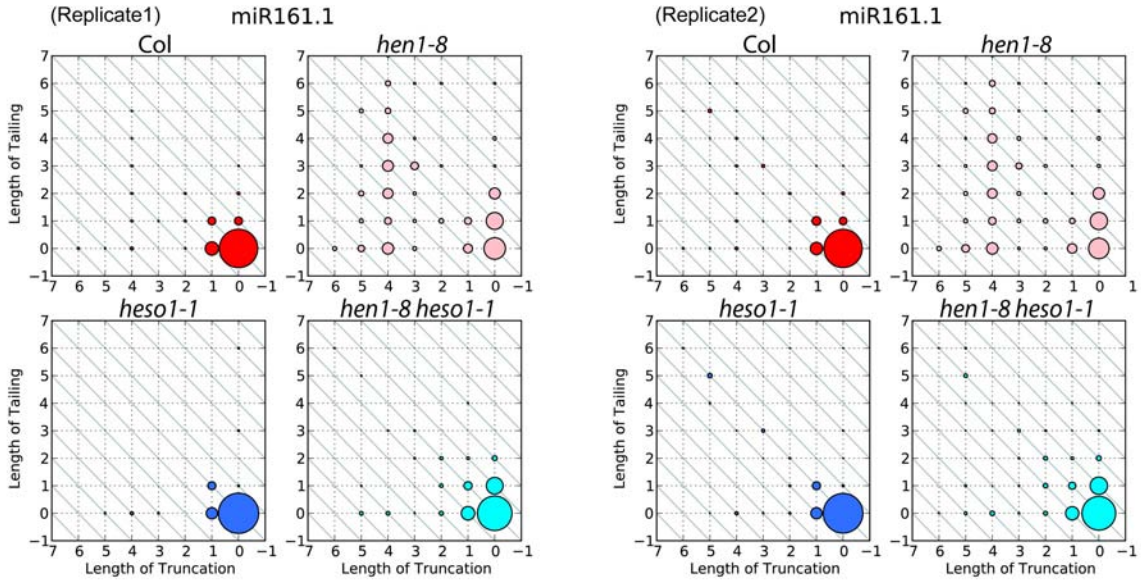


Figure S3-3

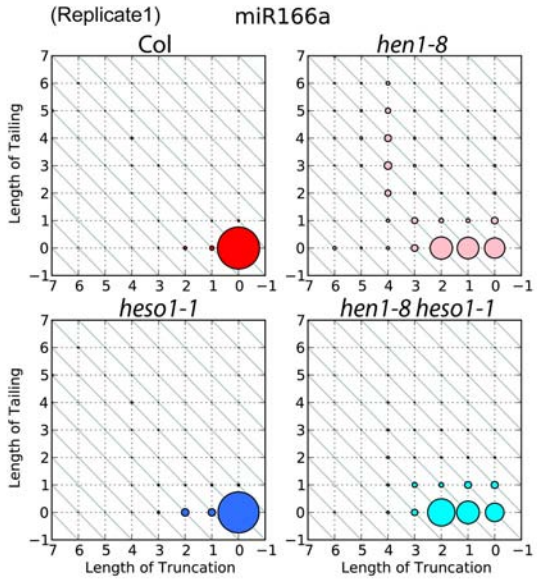


Figure S3-4

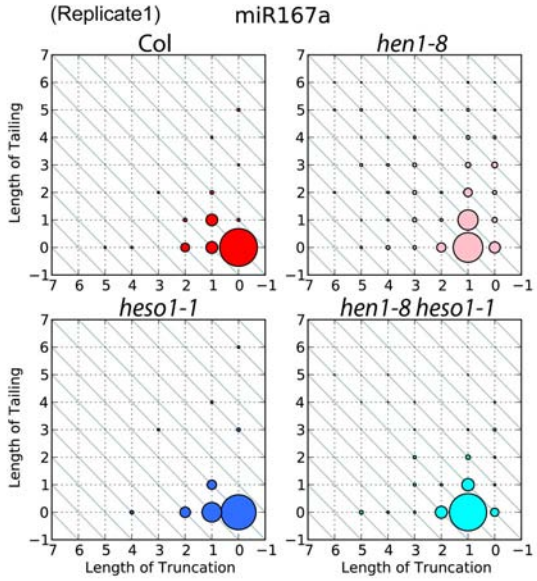


Figure S3-5

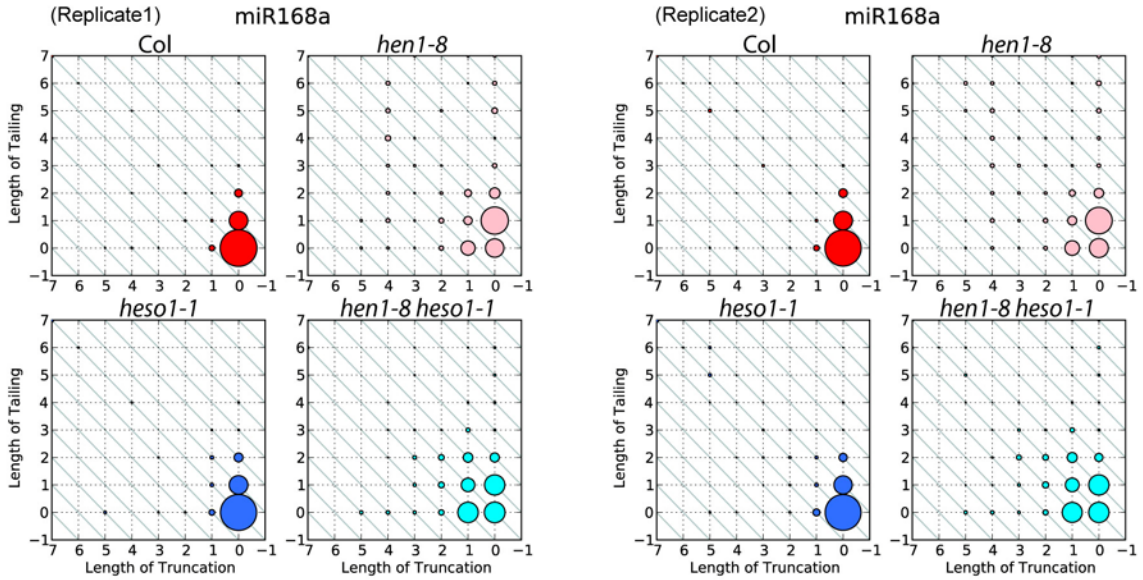


Figure S3-6

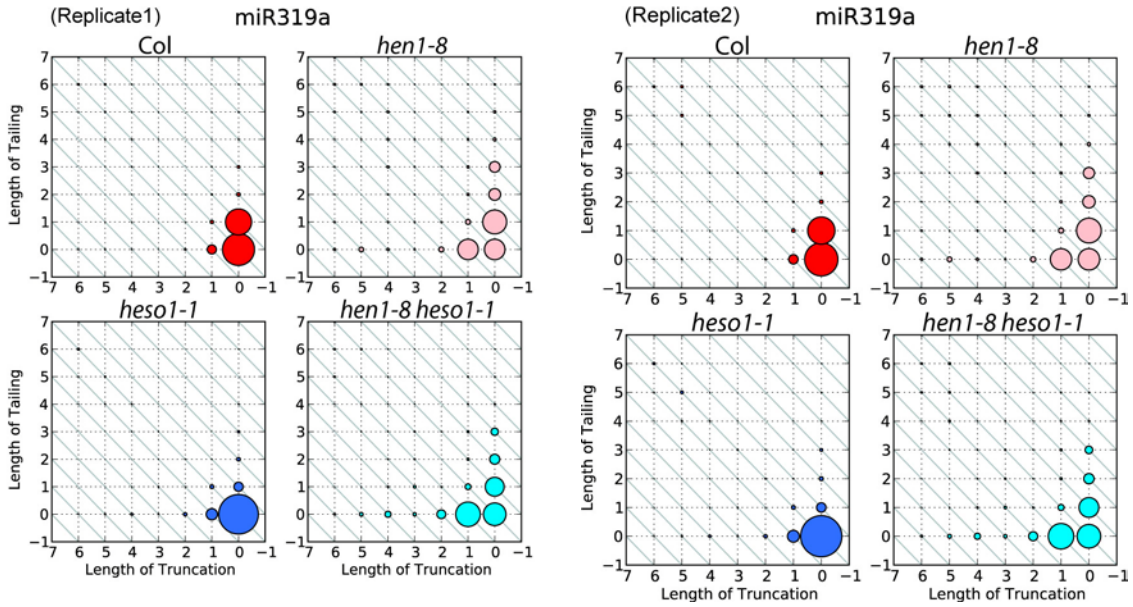


Figure S3-7

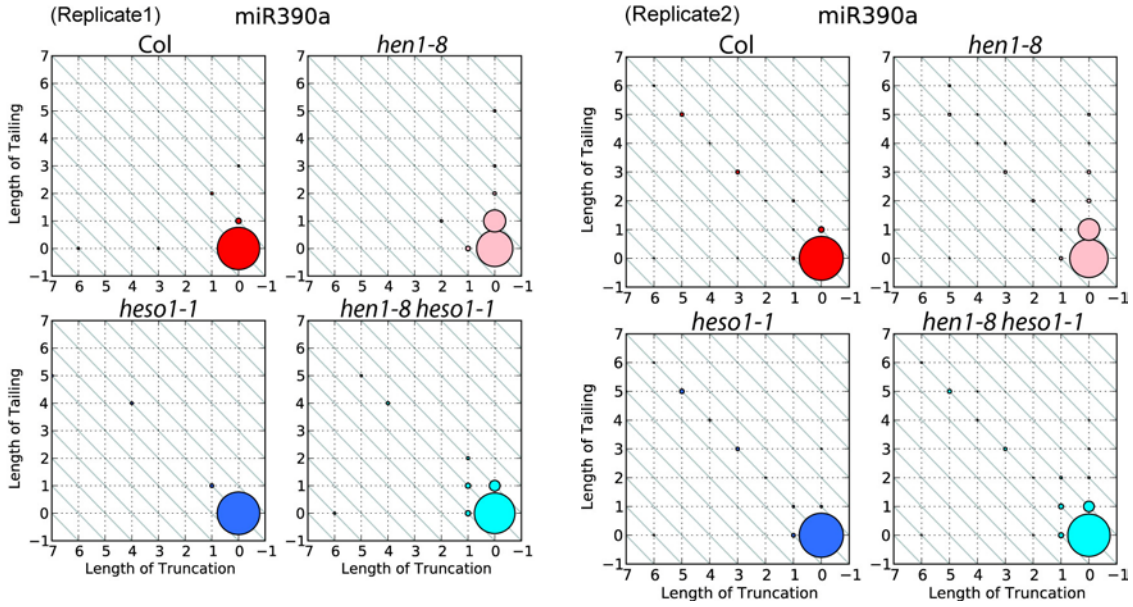


Figure S3-8

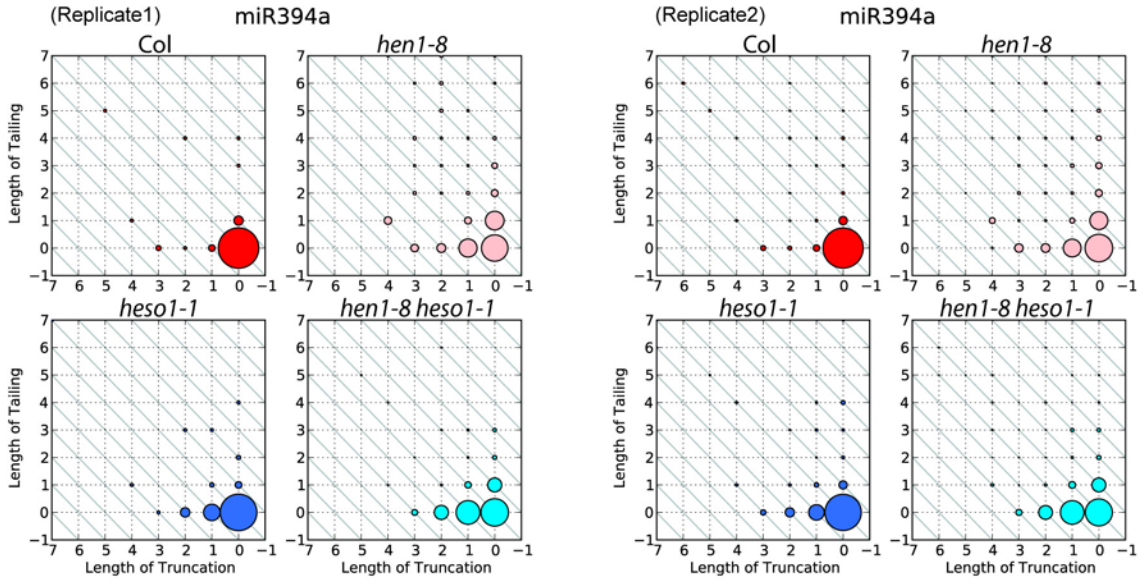


Figure S3-9

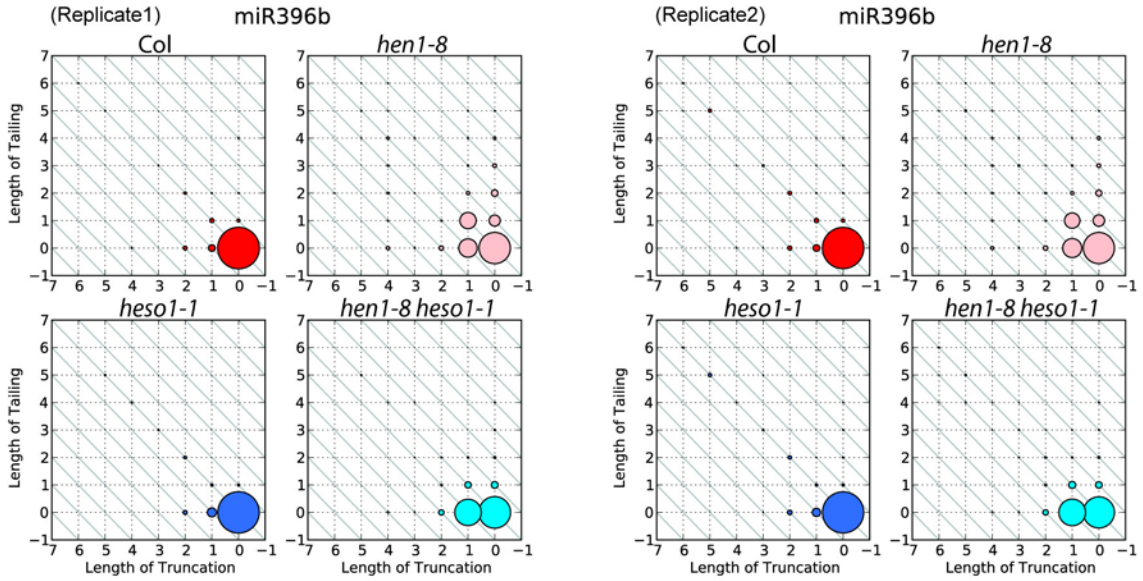


Figure S3-10

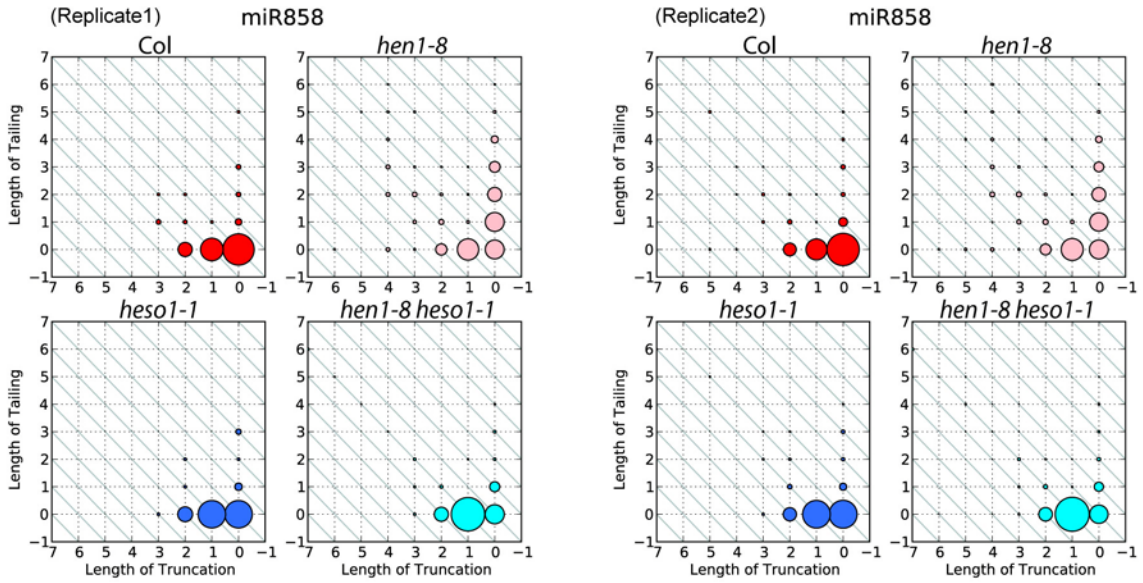


Figure S3-11

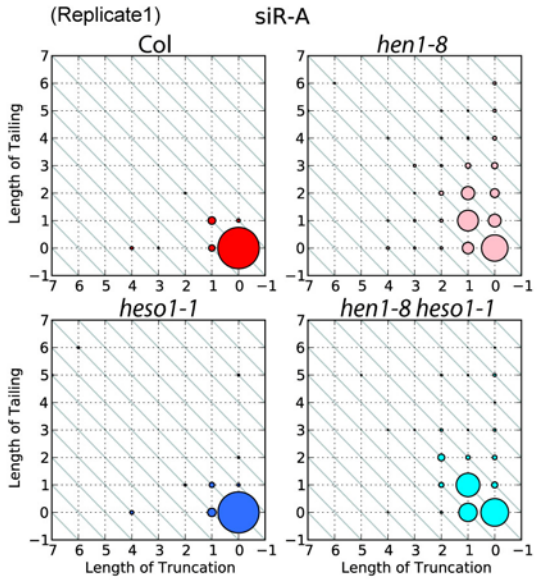
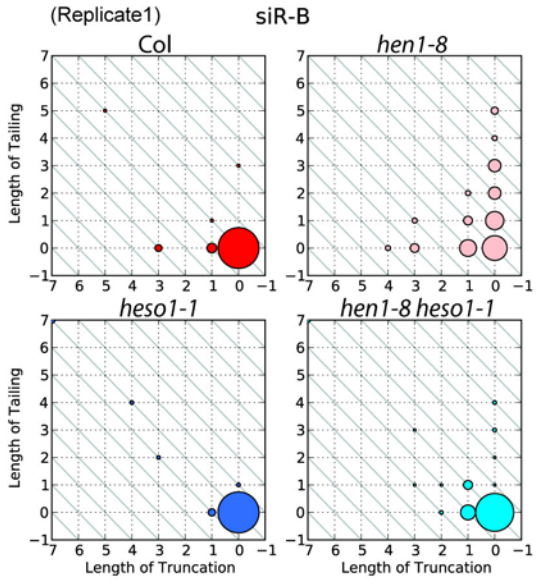


Figure S3-12



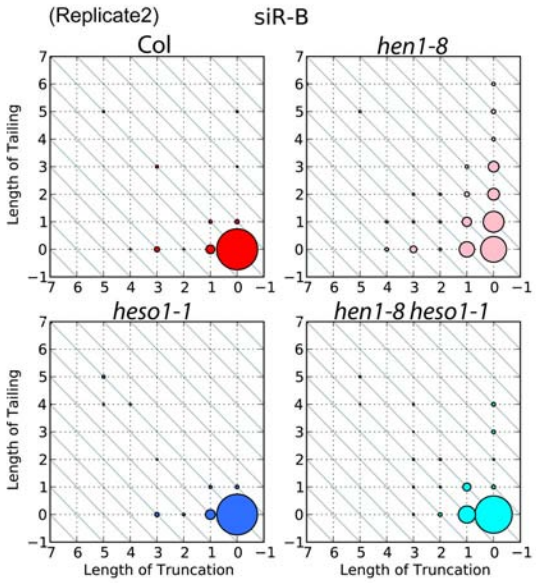


Figure S3-13

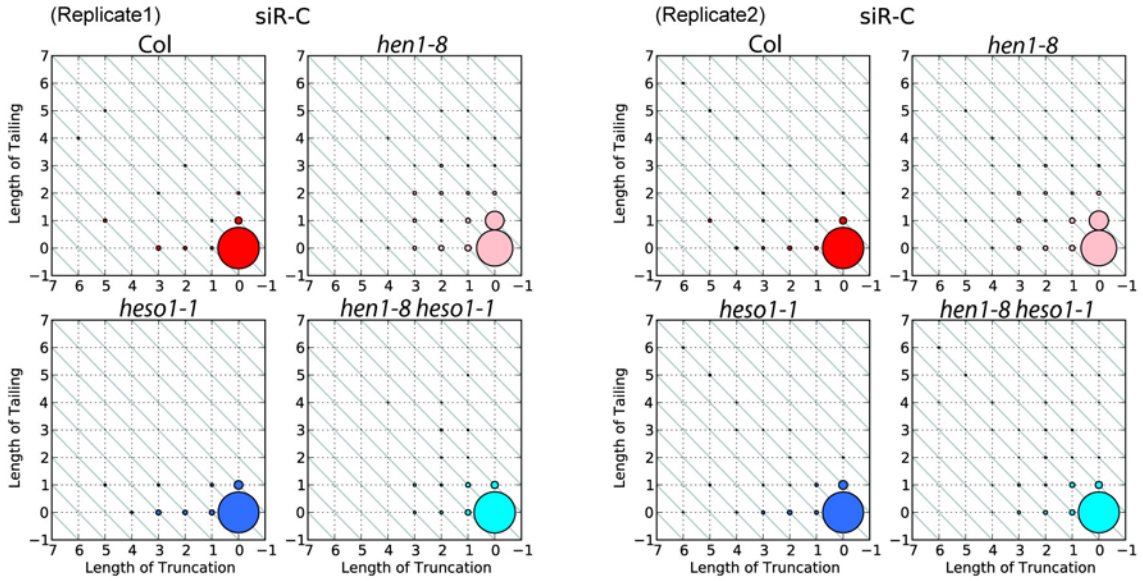


Figure S3. The *heso1-1* mutation reduced 3' tailing of miRNAs and endogenous siRNAs without much effects on 3' truncation in *hen1-8*, Related to Figure 3

The status of 3' truncation and/or 3' tailing for each small RNA is represented by a two dimensional matrix in which the X-axis represents the extents of 3' truncation, the Y-axis represents the extents of 3' tailing, and the sizes of the circles indicate the relative abundance of the small RNA species. In *hen1-8*, both 3' truncation and 3' tailing occurred at a much higher frequency than in wild type. In *hen1-8 heso1-1*, 3' tailing was drastically reduced but 3' truncation was largely unaffected. The sequences of the siRNAs are listed in Table S1. Matrices for seven miRNAs and two siRNAs from the second biological replicate and the same small RNAs from the first biological replicate are shown. In addition, matrices for miR159a, miR166a, miR167a, and siR-A from the first biological replicate are shown (the matrices for these small RNAs from the second biological replicate are shown in Figure 3).

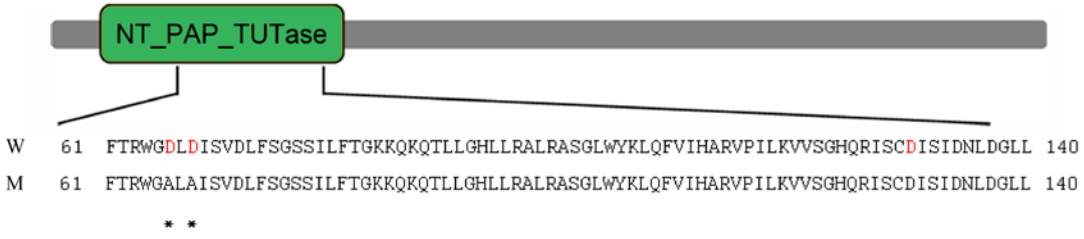


Figure S4. The nucleotidyl transferase (NT) domain in the HESO1 protein, Related to Figure 4

Part of the NT domain sequence is shown below. Three conserved aspartic acids (D66, D68 and D130 in red) form a metal binding triad in the wild-type (W) protein. In the mutated version (M) used for the *in vitro* enzymatic assay, the first two aspartic acid residues were mutated to alanine (marked by *).

Table S1. Sequences of Oligonucleotides Used in This Study

Purpose	Primer name	Oligonucleotide sequence (5'-3')
Genotyping	GK367_GT-F	CTGGTTCTGTGATTGTTAGGTG
Genotyping	GK367_GT-R	GAGACCAACAGCTCCGAGA
Genotyping	Gabi-kat_T-DNA	ATATTGACCATCATACTCATTGC
Genotyping	T13K14-P19	TAGTCTCATCAAGTTATGTCT
Genotyping	T13K14-P20	GCAGAGAAGCGTGTTCAATC
Plasmid construction	at2g39740_SacII	tccccgcggTGAGGATTGAGGTAAGCCA
Plasmid construction	at2g39740_EcoRI_R_Nstop	cggaattcCTGCTCATGTCTCGGTCTC
Plasmid construction	HESO1-EcoRI_F-28b	cggaattcgATGAGTAGAAACCCTTTCC
Plasmid construction	HESO1-EcoRI_R-28b	cggaattcCTACTGCTCATGTCTCGGT
Plasmid construction	HESO1_DADA_F	TCACACGATGGGGAgccTTAgctATCTCTGTTGACTT
Plasmid construction	HESO1_DADA_R	AAGTCAACAGAGATagcTAAggcTCCCCATCGTGTGA
Northern blot	miRNA166_AS	GGGGAATGAAGCCTGGTCCGA
Northern blot	miRNA173_AS	GTGATTTCTCTCTGTAAGCGA
Northern blot	miRNA167_AS	TAGATCATGTTGGCAGTTTCA
Truncation and tailing analysis	siR-A	AACTTGGATACTGTGAATGATGCA
Truncation and tailing analysis	siR-B	ATTGAGGCTGAGTCGAGTAGAGCT
Truncation and tailing analysis	siR-C	AAGAATCAACGAACTGCAGACACC
Enzymatic activity assay	miR173 RNA oligo	UUCGCUUGCAGAGAGAAAUCAC
Real-time PCR Ubiquitin 5	UBIQUITIN5-F	GGTGCTAAGAAGAGGAAGAAT
	UBIQUITIN5-R	CTCCTTCTTTCTGGTAAACGT
Real-time PCR CUC2	CUC2F2	CTCACTCCCACCTCTCCT
	CUC2R2	GAAAAGGGTCAAAGTCAAAC
Real-time PCR SCL6	SCL6_F	GAGGTCATAGAGAGCGACAC
	SCL6_R	GGAATAAGGGTTTAGGGTTT
Real-time PCR SPL10	SPL10_F	TGAGACAAAGCCTACACAGATGGA
	SPL10_R	GATGATGCAACCCGACTTTTTTATG
Real-time PCR PHB	PHB_realtimeF1	ATGAAACATCAACTTCACACTG
	PHB_realtimeR2	CGTTTCCAGCAGGTATCACA
RT-PCR	NTP4-5'RT_F	GGAGCAAGACGATGAGTAGA
RT-PCR	NTP4-5'RT_R	GTAGTAAATGTCCGAGTAAAGT
RT-PCR	ACTIN 8 F	CACATGCTATCCTCCGTCTC
RT-PCR	ACTIN 8 R	CAATGCCTGGACCTGCTT

Supplemental Experimental Procedures

Arabidopsis Strains and Growth Conditions

Plants were grown under long day (16 h light/ 8 h darkness) conditions at 22°C. All *Arabidopsis* strains are in the Columbia background. Seeds of *heso1-1* were obtained from the Gabi-Kat (GK-367H02-017041) collection. To obtain the *hen1-8 heso1-1* double mutant, *hen1-8* was crossed with *heso1-1* and the F2 population was genotyped for *hen1-8* and *heso1-1* mutations. The *hen1-8* mutation was genotyped by PCR using primers T13K14-P19 and T13K14-P20 (Table S1) followed by restriction digestion with *Hpa*I (only the mutant sequence can be digested). *heso1-1* was genotyped by PCR with the primer pair GK367_GT-F/GK367_GT-R (Table S1), which amplifies the wild-type genomic fragment (absent in *heso1-1* homozygous plants), and PCR amplification with the primer pair GK367_GT-R/Gabi-kat_T-DNA (Table S1), which amplifies part of the T-DNA insertion (present in the *heso1-1* homozygous and heterozygous plants).

Plasmid Construction

To construct the *pHESO1:HESO1-GFP* plasmid, the *HESO1* genomic region from a 2kb promoter to the end of the coding region (without the stop codon) was amplified by PCR with primers at2g39740_SacII and at2g39740_EcoRI_R_Nstop (Table S1), and cloned into pJL-blue, a pBluescript-based Gateway entry vector [1], at *Sac*II and *Eco*RI sites. Then the DNA fragment was moved into pMDC107 [2] by the LR Gateway reaction. This plasmid was used to transform *hen1-8 heso1-1* plants by *Agrobacterium* (GV3101) mediated floral dip transformation. To construct the *6xHis-HESO1_WT* plasmid, the *HESO1* coding sequence was amplified using the cDNA clone S67294 (obtained from ABRC) as a template with primers HESO1-EcoRI_F-28b and HESO1-EcoRI_R-28b (Table S1). The amplified fragment was cloned into *pET-28b* at the *Eco*RI site. The orientation of the insertion was confirmed by sequencing.

To construct the *6xHis-HESO1_M* plasmid, the *HESO1* coding sequence with the D66A and D68A mutations was first generated by PCR using primers that incorporated the mutations and clone S67294 as a template. The left fragment (primers: HESO1-EcoRI_F-28b and HESO1_DADA_R) and right fragment (primers: HESO1_DADA_F and HESO1-EcoRI_R-28b) (Table S1) were annealed and used as the template to amplify *HESO1 CDS_M* using HESO1-EcoRI_F-28b and HESO1-EcoRI-28b. This fragment was cloned into *pET28b* at the *Eco*RI site. The orientation of insertion was confirmed by sequencing.

Reverse Transcription and PCR

Reverse transcription and real-time PCR were performed as previously described [3]. *UBIQUITIN5* was used as an internal control for the real-time RT-PCR to examine the expression of miRNA target genes. *ACTIN8* was used as an internal control for the *HESO1* RT-PCR. Primers used are listed in Table S1.

Northern Blotting

RNA isolation and northern blotting to detect small RNAs were performed as described [4]. Total RNAs were extracted from inflorescences using TRIzol reagent (Invitrogen). 5′-end-labeled (32P) antisense DNA oligonucleotides were used to detect miRNAs from total RNAs. Sodium periodate treatment and β-elimination were done as described [5].

Small RNA Library Construction and Sequencing

Small RNAs of a desired size range (15–40 nt) were fractionated from total RNAs by 15% polyacrylamide gel electrophoresis [6]. These small RNAs were sequentially ligated with the 3' and 5' adapters using the Small RNA Sample Preparation Kit (Illumina) according to the manufacturer's instructions. A reverse transcription reaction followed by a low cycle PCR was performed to obtain sufficient amounts of products for deep sequencing. The wild type, *heso1-1*, *hen1-8*, and *hen1-8 heso1-1* libraries were barcoded and sequenced in one channel on an Illumina Hiseq2000. Two biological replicates were performed. The small RNA sequence data are available at GEO under the accession number GSE35479.

Analysis of miRNA and siRNA Tailing and Truncation

Small RNA reads were filtered for quality and separated into different genotypes according to the barcodes using Illumina's built-in pipeline. The raw data are in the process of being submitted to NCBI GEO. Reads that matched to tRNAs and rRNAs were removed. In biological replicate one, the total reads that passed the quality and tRNA/rRNA filters for wild type, *heso1-1*, *hen1-8*, and *hen1-8 heso1-1* were 2525504, 1695482, 2518194, and 2840703, respectively. In biological replicate two, the total reads that passed the filters for the same four genotypes were 7638946, 13470428, 23249325, and 15341388. The small RNA reads were then mapped to the *Arabidopsis* genome using Bowtie [7]. To analyze miRNA 3' derivatives, any small RNA read that could not be perfectly mapped back to the genome was chopped one nucleotide at a time from the 3'end until the remaining sequence was perfectly mapped to the genome. Thus, for any non-genome-matched small RNA read, the sequence was split into two parts: the longest 5' genome-matched component (5GMC) part, and a 3' "tail" part. For any read that could be mapped to the genome without chopping, the 5GMC would be the same as the read and there would be no tail. With all reads processed into the format of 5GMC plus tail (the tail could be null), the 5GMC of each read was compared to all annotated miRNAs in miRBase to ascertain those originating from known miRNAs. miRNAs with a 5GMC frequency of 50 TPM (transcript per million) and above were included in the tailing analyses.

To determine the extent of "tailing" (addition of non-templated 3' nucleotides) and "truncation" (shortening from the 3' end), *Arabidopsis* miRNAs annotated in miRBase v17 were examined. Small RNA reads with the 5GMC portion perfectly aligning to each one of the annotated miRNAs were identified, and the proportion of tailing and truncation was calculated based on read frequencies for the different categories of forms (full-length, tailed only, truncated only, and tailed and truncated) for each miRNA. For siRNAs, we chose a small set of 24 nt siRNAs (see Table S1 for their sequences) that are abundantly represented for the analysis. Small RNA reads matching to these siRNA sequences were separated into 5GMCs and 3' tails as for miRNAs.

Protein Purification and Enzymatic Assays

The *6xHis-HESO1_WT* and *6xHis-HESO1_M* plasmids were transformed into the *E. coli* strain BL21 Star™(DE3) (Invitrogen) for protein expression. The transgenic *E. coli* strains were cultured at 37°C until the OD reached 0.5. IPTG was added to a final concentration of 0.2 mM and the culture was incubated at 16°C overnight. The recombinant proteins were purified using Ni-NTA agarose (Qiagen) under native conditions following the manufacturer's instructions with some modifications. During the wash step, the column was washed in succession with the

following buffers: 10 volume of wash buffer, 10 volume of wash buffer with 50% glycerol and 1% NP-40, 5 volume of wash buffer, 10 volume of wash buffer with 3M Urea and 1% NP-40, 5 volume of wash buffer, 8 volume wash buffer with 100 mM imidazole, and 2 volume of wash buffer with 120 mM imidazole. Then the protein was eluted with elution buffer.

miR173 RNA oligonucleotides (either with a 2'-OH or with a 2'-*O*-methyl group at the 3' terminal nucleotide) were labeled in a 50 μl reaction mixture containing 100 μM miR173, 50 U T4 Polynucleotide Kinase (New England Biolabs), 1X T4 PNK buffer and 5 μl ATP [γ-32P] (3000Ci/mmol 10mCi/ml from PerkinElmer). After incubating at 37°C for 1 hour, the mixture was purified with Illustra MicroSpin™ G-25 Columns (GE Healthcare) according to the manufacturer's instructions.

The HESO1 enzymatic activity assays were conducted in 20 μl reaction mixtures containing 1 μg recombinant wild-type or mutated 6XHis-HESO1, 1 μl labeled miR173, 20 mM Tris-HCl (pH 8.0), 50 mM KCl, 0.7 mM $MnCl_2$, 2.5 mM $MgCl_2$, 0.5 mM DTT, 200 μg/mL BSA, and 0.5 mM of different nucleotide triphosphates. Reactions were incubated at room temperature for 0-40 min, and stopped by adding the formamide small RNA loading dye to the reaction mix. Then, the mix was denatured at 75°C, incubated on ice for 5 min, then the RNAs were separated on a 15% acrylamide urea gel and visualized by autoradiography.

Supplemental References

1. Yant, L., Mathieu, J., Dinh, T.T., Ott, F., Lanz, C., Wollmann, H., Chen, X., and Schmid, M. (2010). Orchestration of the floral transition and floral development in Arabidopsis by the bifunctional transcription factor APETALA2. Plant Cell *22*, 2156-2170.

2. Curtis, M.D., and Grossniklaus, U. (2003). A gateway cloning vector set for high-throughput functional analysis of genes in planta. Plant Physiol *133*, 462-469.

3. Kim, Y.J., Zheng, B., Yu, Y., Won, S.Y., Mo, B., and Chen, X. (2011). The role of Mediator in small and long noncoding RNA production in Arabidopsis thaliana. The EMBO journal *30*, 814-822.

4. Park, W., Li, J., Song, R., Messing, J., and Chen, X. (2002). CARPEL FACTORY, a Dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in Arabidopsis thaliana. Current biology *12*, 1484-1495.

5. Yu, B., Yang, Z., Li, J., Minakhina, S., Yang, M., Padgett, R.W., Steward, R., and Chen, X. (2005). Methylation as a crucial step in plant microRNA biogenesis. Science *307*, 932-935.

6. Lu, C., Meyers, B.C., and Green, P.J. (2007). Construction of small RNA cDNA libraries for deep sequencing. Methods *43*, 110-117.

7. Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol *10*, R25.