DEEP SEQUENCING FROM *hen1* MUTANTS

TO IDENTIFY SMALL RNA 3' MODIFICATIONS

by

Jixian Zhai

A thesis submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Master of Science in Bioinformatics and Computational Biology

Spring 2013

© 2013 Jixian Zhai
All Rights Reserved

DEEP SEQUENCING FROM *hen1* MUTANTS

TO IDENTIFY SMALL RNA 3' MODIFICATIONS

by

Jixian Zhai

Approved: ____

Blake C. Meyers, Ph.D.
Professor in charge of thesis on behalf of the Advisory Committee

Approved: ____

Errol Lloyd, Ph.D.
Chair of the Department of Computer and Information Sciences

Approved: ____

Babatunde A. Ogunnaike, Ph.D.
Interim Dean of College of Engineering

Approved: ____

James G. Richards, Ph.D.
Vice Provost for Graduate and Professional Education

ACKNOWLEDGMENTS

I wish to thank my adviser, Dr. Meyers, and my committee members Dr. Liao Li and Dr. Adam Marsh for their continuous advice, guidance, and academic support during my master study. I would also like to thank Dr. Cathy Wu and Katie Lakofsky in the Bioinformatics program for their great work in carrying out this master program. I must also thank my friends Katie Bi, Yuanlong Tian, Jia Ren and Qili Fei who have supported and helped me throughout my master education.

This manuscript is dedicated to my parents, Faqing Zhai and Yulan Zhao for their unconditional love. (这篇论文献给我的父母，翟发庆和赵玉兰，感谢他们对我一直以来无尽的爱与支持。）

TABLE OF CONTENTS

LIST OF FIGURES v
ABSTRACT vi

Chapter

1 INTRODUCTION 1

1.1 Small RNAs in plants 1
1.2 Discovery of HEN1, and its conserved function in many species 3
1.3 3’ modifications of plant small RNAs: HEN1 and its sidekicks 7

2 METHODS 10

2.1 Overview of the Pipeline 10
2.2 Visualization of the truncation and tailing pattern 13

3 APPLICATIONS 16

3.1 Small RNAs 3’ modifications in *hen1* and *heso1* mutants 16
3.2 Portion of tailed miRNA variants in *hen1* and *heso1* mutants 18

4 CONCLUSION 20

5 REFERENCES 22

Appendix

PUBLICATION RELATED TO THIS WORK 26

LIST OF FIGURES

Figure 1: Possible 3' modifications on small RNAs. Small RNAs may be 3' modified in several ways in plant hen1 mutants 11

Figure 2: Data handling for truncation and tailing analysis.. 13

Figure 3: Matrix for summarizing and visualizing miRNA 3' truncation and tailing from deep sequencing data.. 15

Figure 4: heso1-1 mutation reduces 3' tailing on miRNAs in hen1-8.. 17

Figure 5: The portion of 3' uridylation of miRNA variants and their sequence composition. 19

ABSTRACT

MicroRNAs (miRNAs) function via targeting of messenger RNAs, suppressing protein levels, and playing important roles in biological processes of plants and animals. The pathway for miRNA biogenesis is well established, but less is known about miRNA turnover, largely due to difficulties in capturing miRNAs during the process of decay, in which they are both rare and ephemeral. The HEN1 protein methylates the 3' terminus of small RNAs (small RNAs), protecting them from poly-urydilation and degradation. Recent progress using deep sequencing to study small RNAs in *hen1* reveals the potential for *hen1* to serve as a platform for studies of miRNA turnover, with the sequencing data providing unprecedented precision and detail in the characterization of 3’ modifications.

Chapter 1

INTRODUCTION

1.1 Small RNAs in plants

In plants, microRNAs (miRNAs) [1, 2] and small interfering RNAs (siRNAs) [3, 4] are the two major classes of small RNAs. MicroRNAs in plants function in a wide variety of roles principally in development, but have demonstrated roles in modulation of transcriptional responses to abiotic and biotic stress, and to nutrient deficiencies [5-8]. Small interfering RNAs in plants are predominantly but not exclusively composed of heterochromatic siRNAs, which function in genome defense via the maintenance of heterochromatin, presumably important for the suppression of transposon activity [9].

MiRNAs are processed from long single-stranded transcripts that originate from distinct genomic loci and are transcribed by RNA polymerase II (Pol II) like a typical mRNA. These miRNA precursor transcripts fold into imperfectly-paired, hairpin-like structures and are subsequently processed via a series of nuclear-localized events (in plants, as some processing events in animals occur in the cytoplasm) that lead to the generation of a single, mature small RNA of typically 21 or 22 nt [10, 11]. In an early processing step, a small RNA "duplex" comprised of both the miRNA and

miRNA-star is cleaved from the precursor hairpin by the DICER-LIKE1 (DCL1) protein. The miRNA strand, also known as the "guide" molecule eventually becomes associated with (or "loaded" into) an ARGONAUTE (AGO) protein to form an active RNA-Induced Silencing Complex (RISC). The miRNA-star strand is usually non-functional (hence it is also known as the "passenger" strand), and it is preferentially decayed leading to a substantially lower abundance than the miRNA. In contrast, siRNAs are often not *per se* encoded in the genome but are derived from precursor RNA molecules that have been made double-stranded by RNA-dependent RNA polymerases (RDRs). These dsRNA precursors are processed by DCL proteins into multiple distinct siRNA classes that represent both strands of the original dsRNA [3, 12]. Plant siRNAs come in several forms; for example, heterochromatic siRNAs are produced from RNA polymerase IV-dependent precursors generated from a wide variety of repetitive sequences [13], while trans-acting or phased siRNAs are produced via Pol II from long, non-coding mRNAs or even from protein-coding transcripts [14-16]. The *Arabidopsis* genome includes four *DCL* genes that are responsible for generating distinct classes of small RNAs and of different sizes, ranging from 21- to 24-nt. Mutant analysis has revealed some functional redundancies among the *DCL* genes [17, 18]. Despite the diversification of small RNA types and biogenesis pathway proteins, current data suggest that all plant small RNAs require 3' methylation performed by a single protein known as HEN1. This protein also has a conserved function in the maturation of specific subclasses (namely the "piRNAs") of animal small RNAs. Deep sequencing of small RNAs in mutants of *hen1* can describe

the types of 3' modifications to which small RNAs are subjected in the absence of a 3' methyl group. In this article, we will discuss the use of these deep sequencing data, the types of modifications that are observed, and the insights that may be made from these analyses when performed on *hen1* mutants.

1.2 Discovery of HEN1, and its conserved function in many species

The *HEN1* (*HUA ENHANCER1*) gene was first identified in a screen for Arabidopsis floral development mutants [19]. The phenotype was identified as an enhancer of a floral-organ phenotype in the *hua1-1 hua2-1* double mutant; when the Chen lab segregated out the *hen1* mutant from the *hua* mutant genes, the single gene mutant *hen1* was observed to have a pleiotropic phenotype, including a short stature, delayed flowering, reduced fertility, and multiple inflorescence aberrations [19]. These phenotypes were later revealed to be due to widespread reductions in small RNAs and their activities. In subsequent work by the Chen lab, the HEN1 protein was shown to function as a methyltransferase; the methyltransferase activity is specific to small RNAs, adding a 2'-O-methyl group on the 3' terminal nucleotide of template small RNAs [20, 21]. In more recent work, the crystal structure of HEN1 suggests that it can specifically recognize the small RNA duplex as its template via multiple RNA binding domains and a methyltransferase domain [22]. Thus HEN1 activity is highly specific, but a loss of *HEN1* can have a widespread impact on small RNAs, particularly in

plants in which all known small RNA pathways converge on the methylation activity of HEN1.

MiRNAs examined by RNA gel blot analysis exhibit unusual length heterogeneity in a *hen1* mutant background [23]. Comparisons performed by cloning and sequencing individual miRNAs in wildtype and *hen1-1* backgrounds showed the addition of predominantly uridine (U) nucleotides at the 3' end of some miRNAs in a *hen1-1* mutant [23]. This U "tailing" is presumably a result of deprotection caused by a lack of 3' methylation due to the loss of HEN1 activity. RNA gel blot analysis of miRNAs has demonstrated that miRNAs in plant *hen1* mutants display a "laddering" of length, reflecting the 3' tails, while sizes shorter than the wildtype length are also observed, suggesting that 3' truncation occurs as well (Figure 1A) [23]. As we describe below, our analyses of deep sequencing data have confirmed that both 3' truncation and 3' tailing are common for most miRNAs in plant *hen1* mutants. An additional and likely related observation was that the abundance of most miRNAs is considerably reduced in the *hen1* mutants, with no corresponding decrease in the rate of miRNA biogenesis, implying that in *hen1*, miRNAs have increased rates of degradation [20, 23]. At that point in time, it was unclear whether U tailing is a cause or a consequence of miRNA degradation. The 3' tailing of miRNAs is easily detected by today's "next-generation" sequencing, and thus small RNAs in plant *hen1* mutant are good candidates for molecular characterization using modern techniques.

The *hen1-1* allele in the *Arabidopsis* Landsberg *erecta* (L*er*) ecotype came out of an enhancer screen of the *hua1-1*/*hua2-1* background [19]. It was the first *hen1* mutant to be identified in any species and therefore contributed the name by which even animal *HEN1* genes are now known. In Arabidopsis, a variety of alleles have since been characterized, via both forward and reverse genetics approaches. These include the weaker *hen1-2* allele, also from L*er*, and the identical point mutation in the Columbia (Col) background, isolated from an independent genetic screen and known as *hen1-8* [24]; curiously, the same allele in different backgrounds produces different phenotypes, with a more severe developmental impact (on fertility, for example) in the Col background, as well as greater molecular effects such as a reduction in miRNA levels [24]. In rice, two mutant alleles in the ortholog *WAVY LEAF1* (*WAF1*) of Arabidopsis *HEN1* have been described; the rice *waf1-1* and *waf1-2* mutants each bear a single-base substitution, leading to a premature stop codon in the second exon and a non-functional splicing site of the fourth intron, respectively [25]. While *HEN1* is also single-copy in rice, there are numerous phenotypic differences in the rice *hen1* (= "*waf*") mutants, compared to Arabidopsis, most of which reflect a more severe impact on normal biological functions. For example, the rice *hen1* mutants exhibit seedling lethality and phenotypes reminiscent of tasiRNA defective mutants in rice [25]. Because mutants in the rice ortholog were identified only recently, work on the plant *HEN1* gene has been primarily performed in *Arabidopsis*, however, the characterization of *wavy leaf1* (*waf1*) mutants represents an opportunity to expand our understanding of HEN1 function in monocots.

Besides plants, HEN1 orthologs have also been described and studied in metazoans, in which HEN1 methylates different subsets of small RNAs, primarily including piwi-interacting RNAs (piRNAs) but excluding miRNAs. PiRNAs are a class of animal small RNAs which are abundant in germline cells; they are longer and distinct from miRNAs, and target transposons to effect silencing [26]. Unlike animal miRNAs which are unmethylated (as a reminder, plant miRNAs are methylated), piRNAs are methylated, a modification performed by HEN1. Thus, the mouse *HEN1* homolog (*mHEN1*) is expressed specifically in the testes, where it functions to methylate piRNAs [27-29]. HEN1 substrates in *Drosophila* are determined by AGO binding - AGO1-bound miRNAs remain unmethylated, but AGO2-bound siRNAs and Piwi-bound piRNAs are methylated [30, 31]. This suggests that animal HEN1 acts on AGO-bound single-stranded small RNAs, whereas the plant HEN1 acts on small RNA duplexes before they are incorporated into AGO proteins. The HEN1 homolog in zebrafish has also been studied; as in *Drosophila* and mouse, the zebrafish HEN1 methylates piRNAs in germline cells and is required for oocyte development, and deep sequencing of small RNAs in the zebrafish *hen1* mutant revealed addition of U-tailing on the unmethylated piRNAs [32]. Although miRNAs in animals mostly remain unmethylated, they are fully methylated in plants as a crucial step for stabilization and protection from enzymatic activities such as the 3' truncation and uridylation mentioned above. miRNAs in Arabidopsis *hen1* mutants, without the protection of 3'-methylation, are often uridylated and in some cases truncated.

1.3 3’ modifications of plant small RNAs: HEN1 and its sidekicks

The Chen lab and collaborators have published biochemical data demonstrating the specificity of HEN1 activity. Their initial work in this area demonstrated the methyltransferase activity by which HEN1 transfers a methyl group to the 2’ hydroxyl of the 3’ terminal nucleotide of a small RNA [20, 21]. They demonstrated that both strands of either a miRNA/miRNA* duplex siRNA/siRNA* duplex are methylated *in vitro*, and HEN1 has a preference for 21–24 nt RNA duplexes with a two-nucleotide overhang typical of DICER cleavage [21]. HEN1 in plants possesses double-stranded RNA binding domains (dsRBDs), domains missing in animal orthologs, and structural data demonstrated that plant HEN1 uses the dsRBDs to recognize ~16 bps of its duplexed small RNA substrate, measuring the length of the substrate but transferring the methyl group to the duplex in a non-sequence-specific manner [22]. The transfer of the methyl group to the small RNA duplex occurs via a novel mechanism dependent on Mg^{2+}, coordinated between four invariant residues in the active site of the methyltransferase domain and the 2′ and 3′ hydroxyls on the 3′-terminal nucleotide of the template [22]. Analogous biochemical work on the mouse HEN1 protein has demonstrated that its activity is specific for single-stranded piRNAs [27], consistent with the lack of a dsRBD in the animal orthologs identified to date. Therefore, HEN1 proteins function to add methyl groups to the 3’ end of small RNAs, with no sequence specificity for the substrate yet

demonstrated, other than a preference for either ssRNA (in animals) or dsRNA (in plants).

Other than HEN1, several proteins have been identified for their role in modifying the 3’ end of plant small RNAs. These proteins, functioning as “sidekicks” to HEN1 in 3’ modifications, use unmethylated small RNAs as substrates, presumably those either missed by HEN1 or demethylated via natural processes. Work from the Chen lab in Arabidopsis has identified a family of SDN (SMALL RNA DEGRADING NUCLEASE) proteins that degrade mature miRNAs [33]. While it’s still unclear which proteins are responsible for the 3’ truncation that occurs at a higher rate in *hen1* mutants (Figure 1A), the SDN proteins are candidates for this activity. However, *in vitro* analysis of the SDNs suggests that they degrade miRNAs 3’ to 5’ leaving a small fragment of eight or nine nucleotides in length, and this is quite distinct from the typical 1 or 2 nt shorter size observed for truncated miRNAs in RNA gel blots from *hen1* mutants [20, 23]. More recent work from Chen and Yu labs identified a nucleotidyl transferase, *HEN1 SUPPRESSOR1* (*HESO1*), which is responsible for the addition of the uridine-rich tail to the end of the miRNAs that are not protected by methylation [34, 35]. Other genes in the same family that include poly(U) polymerases were previously shown in animals to modify and regulate the stability of the pre-miRNA (cleaved hairpin) of *let-7* [36]. In the green alga *Chlamydomonas,* it was shown that uridylation added by nucleotidyltransferase MUT68 can promote the degradation of miRNAs and siRNAs [37], and in *C. elegans* CDE-1 uridylates 22G-

RNAs in germline to promote their degradation and thus prevent over accumulation [38]. The methylation of small RNAs by HEN1 blocks both 3’ truncation and 3’ tailing, suggesting that either or both of these 3’ modifications may be important marks for RNA destabilization.

Chapter 2

METHODS

2.1 Overview of the Pipeline

Nearly all studies of 3’ modifications made to miRNAs in *hen1* mutants were performed with RNA gel blots (“northerns”), which can only show the changes in miRNA length and relative abundance, but cannot identify precisely any changes in miRNA sequence composition. In earlier work, cloning followed by Sanger sequencing of miR173 and miR167 in *hen1-1* revealed 3’ truncation as well as one to five nucleotides of uridylation at the 3’ end of miRNAs, consistent with observations on the blots (Figure 1A) [23]. With the tremendous advances in high-throughput sequencing technologies, it is possible to systematically study small RNA populations in *hen1* mutants with unprecedented depth. We have been performing such studies, and while the results are both too extensive to report here and as yet incomplete, the methodologies that we have devised may have utility for the analysis of mutants of *hen1* and the “sidekick” genes that are also involved in 3’ small RNA modifications.

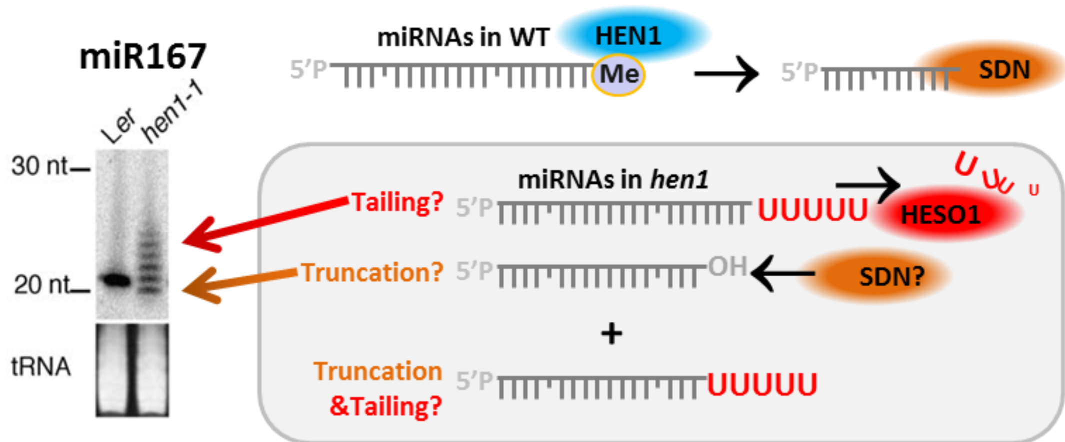

Figure 1 Possible 3' modifications on small RNAs. Small RNAs may be 3' modified in several ways in plant hen1 mutants. The blot on the left shows the laddering effect observed for miR167 in an Arabidopsis hen1-1 mutant versus a Landsberg erecta (Ler) control; reprinted from Li et al. (2005), Current Biology, Vol. 15, 1501-1507 (also cited in the bibliography); copyright 2005, with permission from Elsevier. At top right, in wildtype plants, miRNAs are 3' methylated by HEN1, but may be decayed by the 3' to 5' exonuclease, SDN1/2. In the lower box at right, in a hen1 mutant, miRNAs may be 3' tailed by HESO1, 3' truncated by an as-yet unknown protein (perhaps one of the SDN family of nucleases), or may have some combination of both 3' truncation and 3' tailing.

Our bioinformatics and visualization pipeline assists in the characterization of 3'-end modifications of miRNAs detected using deep sequencing data, typically from an Illumina sequencing instrument. The first step in our process is to identify non-genome-matched small RNA sequences, as most 3' "tailed" small RNAs will no longer match the genome (unless there is a chance identity between the tail and the adjacent nucleotides in the precursor) (Figure 1B). For small RNA reads that were not genome-matched, one nucleotide was chopped off from the 3'end in successive rounds until the remaining 5' sequence can be mapped to the *Arabidopsis* genome with no mismatches (Figure 1B). Thus, any initially non-genome-matched small RNA read could be split into two parts: the longest 5' genome-matched component (5GMC), and a 3' "tail". The tail is nonexistent for reads that matched the genome perfectly in the

first round of analysis. For subsequent analyses, we focus on miRNAs because plant heterochromatic siRNAs are both too variable in sequence (due to their typical origins from repetitive sequences) and too weakly abundant. With all reads processed into the format of 5GMC plus tail, we can also determine the degree of 3' truncation relative to annotated, mature miRNAs sequences, by comparison of the 5GMC to all annotated miRNAs in miRBase for the species of interest [39]. This then determines the extent and composition of the tail (additional, non-canonical 3' nucleotides) and truncation (shortening of the miRNA from the 3' end). The sequence data provide the exact composition of the truncated miRNA and the added 3' tails.

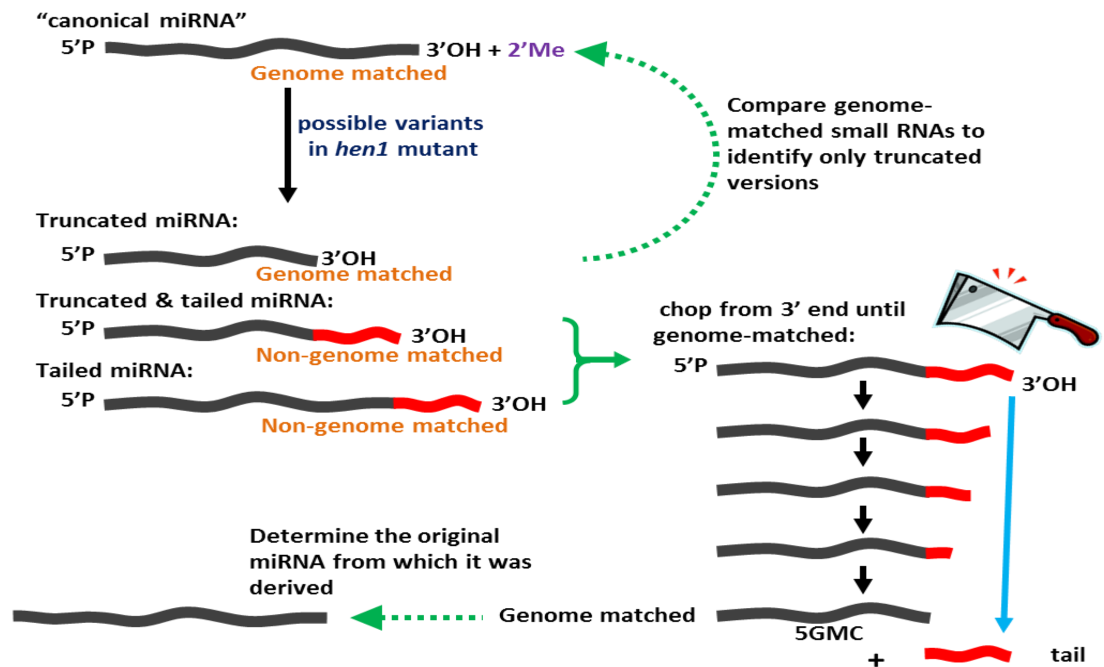

Figure 2 Data handling for truncation and tailing analysis. For small RNA reads that are not genome-matched, one nucleotide is chopped off from the 3' end, in successive rounds until the remaining 5' sequence perfectly maps to genome. Thus, any non-genome-matched sRNA read could be split into two parts: the longest 5' genome-matched component (5GMC), and a 3' "tail". With all reads processed into the format of the 5GMC plus tail, the 5GMC of each read is aligned to annotated miRNAs in miRBase for the genome of origin. This then determines the extent of "tailing" (additional of non-conventional 3' nucleotides) and "truncation" (shortening of the miRNA from the 3' end), as diagrammed in panel A. Each miRNA in an organism could have a different, distinguishable pattern of 3' modifications.

2.2 Visualization of the truncation and tailing pattern

We have also devised a graphic view that simultaneously displays the degree of both truncation and tailing for specific small RNAs (usually individual miRNAs) (Figure 2). In this image, the wildtype or canonical miRNA length is assigned to the position in the lower right of a grid. Using the 5GMC and tailing lengths determined

as described above, miRNA derivatives of identical length can be assigned to different positions in the grid based on the 5GMC value (x-axis, indicating truncation) and the tail length (y-axis). The size of the spot at each position in the grid corresponds to the proportion of total reads for a given small RNA that had that amount of truncation and tailing, regardless of the nucleotide composition of the tail. Any spots on the same diagonal share the same final length and would be indistinguishable on an RNA gel blot, but based on this plot, would have different 3' ends. We provide several examples in Figure 2. For example, Arabidopsis miR166 is significantly truncated and mildly tailed in the *hen1-8* mutant, with a small degree of truncation and tailing visible even in wildtype plants (Figure 2, on left). In *Drosophila* and zebrafish, there is no obvious difference for miRNAs in a mutant *hen1* background, presumably because miRNAs are not methylated in animals; however, we were able to detect a small amount of 3' variation in wildtype animals (Figure 2, middle and right). This could reflect variation in processing or could reflect post-transcriptional 3' modifications resulting from processes similar to those that we have described plants.

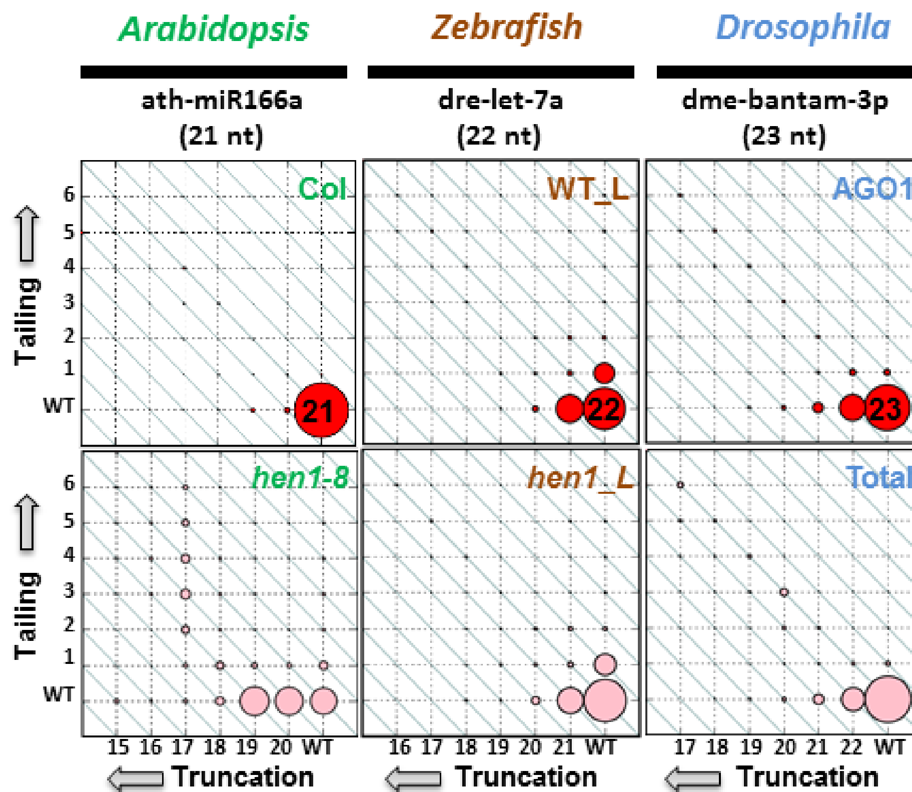

Figure 3 Matrix for summarizing and visualizing miRNA 3' truncation and tailing from deep sequencing data.For each matrix, the x-axis represents the length of the 5' genome-matched component (5GMC) of a particular miRNA related sequence; y-axis represents the length of the "tail" added to the 5GMC, mostly through uridylation. Annotated miRNA sequences are considered to have no truncation or tail, therefore are positioned at (WT, WT) in the lower right corner of each matrix. In this figure, we show the results for a single miRNA (named at the top) from two libraries in each of three organisms, as labeled. The canonical miRNA size is indicated in the spot at the lower right of the top row of panels. In the lower row: for Arabidopsis and zebrafish, we analyzed a hen1 mutant, whereas in Drosophila, we compared AGO1-IPed small RNAs versus total small RNAs. Different colored spots in the top and bottom sets of panels represent different libraries. The data are interpreted as described in the main text. The small RNA sequencing data used in this analysis are described in previous studies. The GEO accession numbers for the libraries we analyzed are GSE35479 (Arabidopsis), GSE33582 (zebrafish) and GSE18806 (Drosophila).

Chapter 3

APPLICATIONS

3.1 Small RNAs 3' modifications in *hen1* and *heso1* mutants

Recent publications from Chen and Yu labs identified HESO1, a poly-uridine polymerase (PUP), that adds the uridine-rich tail to the 3' end of miRNAs that are not methylated [34, 35]. As a co-author in the work from Chen lab [35], I applied our pipeline for miRNA truncation and tailing described above to analyze the small RNA high-throughput sequencing data from wildtype control (Col), hen1-8, heso1-1, and hen1-8/heso1-1 double mutants.

We found that the tailing of miRNAs, which are abundant in the *hen1-8* mutant, is dramatically reduced in the *heso1-1/hen1-8* double mutant, proving that the role of HESO1 in uridylating miRNAs in the absence of methylation.

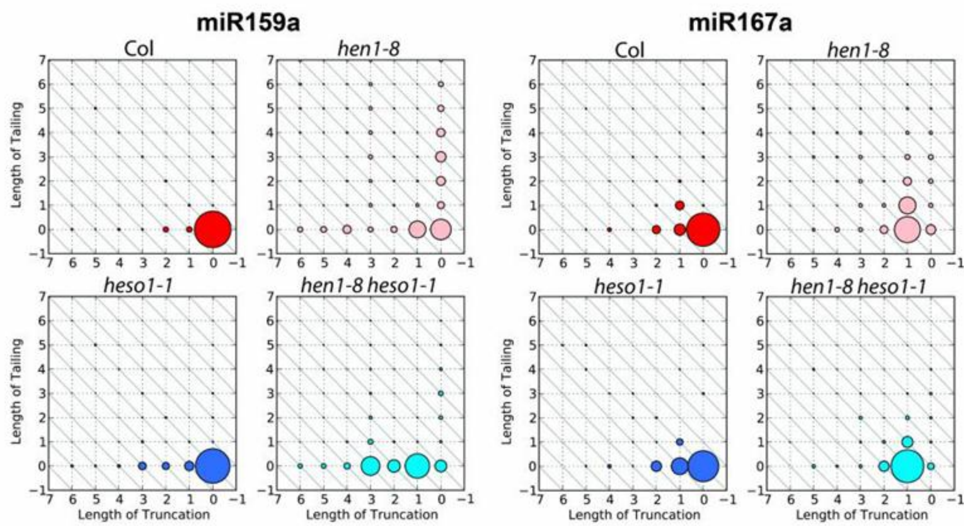

Figure 4 heso1-1 mutation reduces 3’ tailing on miRNAs in hen1-8. From the matrix, both miR159a and miR167a are much more truncated and tailed than in wild type, while in hen1-8/heso1-1 double mutant, 3’ tailing was drastically reduced because HESO1 is the enzyme responsible for uridylation. This figure is published in a paper we co-authored with Chen lab [35].

3.2 Portion of tailed miRNA variants in *hen1* and *heso1* mutants

Because our pipeline can identify all possible variants of a particular miRNA or miRNA family, we can analyze miRNA variants separately based on their type of modifications. And this has proven to be useful for classifying miRNA variants in the heso1 and hen1 data that we collaborated with the Chen lab [35]. In this study, 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). Then, the tailed portion (class 2 plus class 4) were calculated to reflect the role of HESO1 in the tailing of miRNAs in *hen1* mutant. (Figure 5A)

In addition to the sum abundance of all tailed miRNAs, our pipeline can look even further into the sequence composition of the tails. Consistent with previous results that miRNAs are primarily uridylated in *hen1* [23], our results confirm this and showed that the U-rich tailing is reduced in length in the *heso1/hen1* double mutant [35] (Figure 5B).

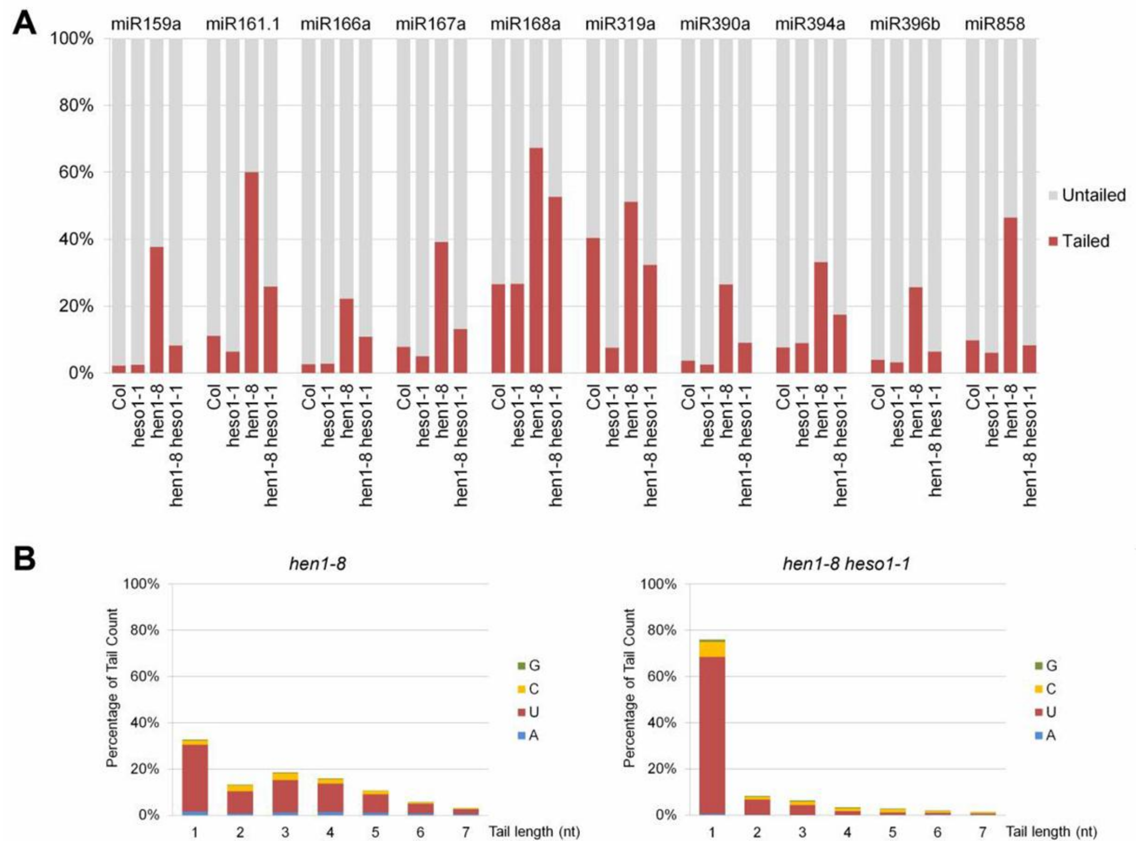

Figure 5 The portion of 3’ uridylation of miRNA variants and their sequence composition. (A) The proportion of tailed and untailed reads (regardless of truncation) for a selection of miRNAs in hen1-8 and hen1-8 heso1-1 double mutant. (B) Sequence composition of “tails” in miR166a variants. Both with predominantly U-rich tailing, tails in heso1/hen1 are significantly shorter than in hen1-8 single mutant. This figure is published in a paper we co-authored with Chen lab [35].

Chapter 4

CONCLUSION

The combination of these analysis and visualization tools will enable dissection of small RNA sequencing data to identify numerous attributes of the post-transcriptional modifications that occur to small RNAs. For example, it will be possible to compare the degree of truncation and tailing across miRNA families within a single organism to assess variability in 3' modifications; these may reflect sequence-specific effects or other factors that have yet to be described. As an example, it's recently been published that AGO10 in Arabidopsis specifically binds miR166 and promotes the degradation of this miRNA [40, 41], so it will be interesting to apply these tools to identify 3' modifications that may be promoted by AGO10. Given that *hen1* mutants are now available for rice, cross-species comparisons of 3' modifications for orthologous miRNAs may also identify conserved or divergent decay destinies for specific miRNAs. One advantage of sequencing-based approaches for studying post-transcriptional modifications to small RNAs is that it is an open-ended analysis which can test all known miRNAs in a single sequencing run. Furthermore, the data facilitate the analysis of other types of small RNAs, including heterochromatic siRNAs and trans-acting siRNAs, although alternative analysis approaches may be required to characterize the truncation and tails of such

heterogeneous populations of small RNAs. We should also point out that there are a number of challenges or limitations to the approaches that we have described. These analyses depend on the accuracy of the miRNA annotation, and recent global analyses have demonstrated that many miRBase-annotated miRNAs have siRNA-like qualities that may reflect older annotations that predate more strict, modern standards [8]. Biological factors can also confound the analyses, including (1) inconsistency in processing of a single precursor (optimal analyses require just one predominant miRNA species per precursor), (2) complex families with many loci that give rise to the same mature miRNAs, (3) weakly abundant miRNAs for which insufficient reads are available.

In conclusion, we believe that these are exciting times in which to study the post-transcriptional events that regulate small RNA function and stability. Numerous plant mutants have been described recently, led by the highly productive forward- and reverse-genetics approaches applied in the Chen lab that have identified a handful of genes critical to 3' modifications of small RNAs. The biochemistry of these genes remains to be worked out, but deep sequencing and genome-wide approaches comprise a powerful and sensitive read-out for the activity of these genes. Future rounds of analyses will likely identify determinants that distinguish the varied activity of these proteins on different small RNAs, while cross-species and even cross-kingdom comparative analyses will provide insights into the biology of these processes.

Chapter 5

REFERENCES

1. Jones-Rhoades, M.W., D.P. Bartel, and B. Bartel, *MicroRNAS and their regulatory roles in plants.* Annu Rev Plant Biol, 2006. **57**: p. 19-53.

2. Reinhart, B.J., et al., *MicroRNAs in plants.* Genes Dev., 2002. **16**(13): p. 1616-1626.

3. Chapman, E.J. and J.C. Carrington, *Specialization and evolution of endogenous small RNA pathways.* Nature Reviews Genetics, 2007. **8**(11): p. 884-96.

4. Hamilton, A.J. and D.C. Baulcombe, *A species of small antisense RNA in posttranscriptional gene silencing in plants.* Science, 1999. **286**(5441): p. 950-2.

5. Rubio-Somoza, I., et al., *Regulation and functional specialization of small RNA-target nodes during plant development.* Curr Opin Plant Biol, 2009. **12**(5): p. 622-7.

6. Sunkar, R., et al., *Small RNAs as big players in plant abiotic stress responses and nutrient deprivation.* Trends Plant Sci, 2007. **12**(7): p. 301-9.

7. Chen, X., *MicroRNA biogenesis and function in plants.* FEBS Lett, 2005. **579**(26): p. 5923-31.

8. Jeong, D.H., et al., *Massive analysis of rice small RNAs: mechanistic implications of regulated microRNAs and variants for differential target RNA cleavage.* Plant Cell, 2011. **23**(12): p. 4185-207.

9. Volpe, T. and R.A. Martienssen, *RNA interference and heterochromatin assembly.* Cold Spring Harb Perspect Biol, 2011. **3**(9): p. a003731.

10. Voinnet, O., *Origin, biogenesis, and activity of plant microRNAs.* Cell, 2009. **136**(4): p. 669-87.

11. Bartel, D.P., *MicroRNAs: genomics, biogenesis, mechanism, and function.* Cell, 2004. **116**(2): p. 281-97.

12. Baulcombe, D., *RNA silencing in plants.* Nature, 2004. **431**(7006): p. 356-63.

13. Wierzbicki, A.T., J.R. Haag, and C.S. Pikaard, *Noncoding Transcription by RNA Polymerase Pol IVb/Pol V Mediates Transcriptional Silencing of Overlapping and Adjacent Genes.* Cell, 2008. **135**(4): p. 635-48.

14. Montgomery, T.A., et al., *Specificity of ARGONAUTE7-miR390 interaction and dual functionality in TAS3 trans-acting siRNA formation.* Cell, 2008. **133**(1): p. 128-41.

15. Cuperus, J.T., et al., *Unique functionality of 22-nt miRNAs in triggering RDR6-dependent siRNA biogenesis from target transcripts in Arabidopsis.* Nat Struct Mol Biol, 2010. **17**(8): p. 997-1003.

16. Zhai, J., et al., *MicroRNAs as master regulators of the plant NB-LRR defense gene family via the production of phased, trans-acting siRNAs.* Genes Dev, 2011. **25**(23): p. 2540-53.

17. Deleris, A., et al., *Hierarchical action and inhibition of plant Dicer-like proteins in antiviral defense.* Science, 2006. **313**(5783): p. 68-71.

18. Henderson, I.R., et al., *Dissecting Arabidopsis thaliana DICER function in small RNA processing, gene silencing and DNA methylation patterning.* Nat Genet, 2006. **38**(6): p. 721-5.

19. Chen, X., et al., *HEN1 functions pleiotropically in Arabidopsis development and acts in C function in the flower.* Development, 2002. **129**(5): p. 1085-94.

20. Yu, B., et al., *Methylation as a crucial step in plant microRNA biogenesis.* Science, 2005. **307**(5711): p. 932-5.

21. Yang, Z., et al., *HEN1 recognizes 21-24 nt small RNA duplexes and deposits a methyl group onto the 2' OH of the 3' terminal nucleotide.* Nucleic Acids Res, 2006. **34**(2): p. 667-75.

22. Huang, Y., et al., *Structural insights into mechanisms of the small RNA methyltransferase HEN1.* Nature, 2009. **461**(7265): p. 823-7.

23. Li, J., et al., *Methylation protects miRNAs and siRNAs from a 3'-end uridylation activity in Arabidopsis.* Curr Biol, 2005. **15**(16): p. 1501-7.

24. Yu, B., et al., *siRNAs compete with miRNAs for methylation by HEN1 in Arabidopsis.* Nucleic Acids Res, 2010. **38**(17): p. 5844-50.

25. Abe, M., et al., *WAVY LEAF1, an ortholog of Arabidopsis HEN1, regulates shoot development by maintaining MicroRNA and trans-acting small interfering RNA accumulation in rice.* Plant Physiol, 2010. **154**(3): p. 1335-46.

26. Juliano, C., J. Wang, and H. Lin, *Uniting germline and stem cells: the function of Piwi proteins and the piRNA pathway in diverse organisms.* Annu Rev Genet, 2011. **45**: p. 447-69.

27. Kirino, Y. and Z. Mourelatos, *The mouse homolog of HEN1 is a potential methylase for Piwi-interacting RNAs.* Rna, 2007. **13**(9): p. 1397-401.

28. Kirino, Y. and Z. Mourelatos, *Mouse Piwi-interacting RNAs are 2'-O-methylated at their 3' termini.* Nat Struct Mol Biol, 2007. **14**(4): p. 347-8.

29. Ohara, T., et al., *The 3' termini of mouse Piwi-interacting RNAs are 2'-O-methylated.* Nat Struct Mol Biol, 2007. **14**(4): p. 349-50.

30. Iwasaki, S., T. Kawamata, and Y. Tomari, *Drosophila argonaute1 and argonaute2 employ distinct mechanisms for translational repression.* Mol Cell, 2009. **34**(1): p. 58-67.

31. Horwich, M.D., et al., *The Drosophila RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC.* Curr Biol, 2007. **17**(14): p. 1265-72.

32. Kamminga, L.M., et al., *Hen1 is required for oocyte development and piRNA stability in zebrafish.* Embo J, 2010. **29**(21): p. 3688-700.

33. Ramachandran, V. and X. Chen, *Degradation of microRNAs by a family of exoribonucleases in Arabidopsis.* Science, 2008. **321**(5895): p. 1490-2.

34. Ren, G., X. Chen, and B. Yu, *Uridylation of miRNAs by hen1 suppressor1 in Arabidopsis.* Curr Biol, 2012. **22**(8): p. 695-700.

35. Zhao, Y., et al., *The Arabidopsis nucleotidyl transferase HESO1 uridylates unmethylated small RNAs to trigger their degradation.* Curr Biol, 2012. **22**(8): p. 689-94.

36. Lehrbach, N.J., et al., *LIN-28 and the poly(U) polymerase PUP-2 regulate let-7 microRNA processing in Caenorhabditis elegans.* Nat Struct Mol Biol, 2009. **16**(10): p. 1016-20.

37. Ibrahim, F., et al., *Uridylation of mature miRNAs and siRNAs by the MUT68 nucleotidyltransferase promotes their degradation in Chlamydomonas.* Proc Natl Acad Sci U S A, 2010. **107**(8): p. 3906-11.

38. van Wolfswinkel, J.C., et al., *CDE-1 affects chromosome segregation through uridylation of CSR-1-bound siRNAs.* Cell, 2009. **139**(1): p. 135-48.

39. Griffiths-Jones, S., et al., *miRBase: tools for microRNA genomics.* Nucleic Acids Res, 2008. **36**(Database issue): p. D154-8.

40. Zhu, H., et al., *Arabidopsis Argonaute10 specifically sequesters miR166/165 to regulate shoot apical meristem development.* Cell, 2011. **145**(2): p. 242-56.

41. Ji, L., et al., *ARGONAUTE10 and ARGONAUTE1 regulate the termination of floral stem cells through two microRNAs in Arabidopsis.* PLoS Genet, 2011. **7**(3): p. e1001358.

Appendix

PUBLICATION RELATED TO THIS WORK

Zhai J and Meyers BC (2013). Deep Sequencing from hen1 Mutants to Identify Small RNA 3' Modifications. Cold Spring Harb Symp Quant Biol.

Zhao Y, Yu Y, **Zhai J**, Ramachandran V, Dinh T, Meyers BC, Mo B, Chen X (2012) HESO1, a nucleotidyl transferase in Arabidopsis, uridylates unmethylated miRNAs and siRNAs to trigger their degradation. *Current Biology.* 22(8):689-94