TRUEE; a bioinformatic pipeline to define the functional microRNA targetome of plants

Gigi Yokchi Wong

June 2022

A thesis submitted for the degree of Doctor of Philosophy of

The Australian National University

Statement of Authorship

The research carried out in this thesis was conducted at The Australian National University between February 2016 and June 2022. I hereby declare that I am the sole author of this thesis. This thesis contains no material previously published except where due acknowledgement has been made and is not being submitted elsewhere for the fulfilment of any other qualification.

Gigi Wong

27th June 2022

Publications and presentations directly arising from this thesis

Chapter 2 has been published:

Wong, G. Y., & Millar, A. A. (2022). TRUEE; a bioinformatic pipeline to define the functional microRNA targetome of Arabidopsis. *The Plant journal: for cell and molecular biology*, *110*(5), 1476–1492. https://doi.org/10.1111/tpj.15751

Figure 3.6 in Chapter 3 has been published in:

Millar, A. A., Lohe, A., & **Wong, G**. (2019). Biology and Function of miR159 in Plants. *Plants*, *8*(8), 255. https://doi.org/10.3390/plants8080255

Parts of Chapter 2 was presented at Bioinfosummer 2020, Australian Mathematical Sciences Institute. “Defining a microRNA targetome in plants: How many genes are regulated by miRNA in Arabidopsis thaliana?”

Acknowledgements

I have been immensely blessed to have had the opportunity to do a PhD in the Millar lab. It has been a long and tough journey but I am so thankful to have run it with the people around me. Tony, you have been such an amazing supervisor. Thank you for teaching me how to think scientifically and for your support throughout the whole journey. You challenged and stretched me which has grown me to be the researcher I am today. I really enjoyed our meetings and talking about innovative directions I could take my project and you inspired and motivated me with your passion. Not just that, but I admire your character too. I really appreciate how you care and support your students and have our best interests in mind (all your students think you're awesome!). Thank you also for supporting me when I struggled with my mental health and for being patient with me despite how often I was grumpy and had a bad attitude. Thank you for lifting me up when I was down and for believing in me even when I didn't. You always saw the positive in me in both my work and in me as a person. I definitely would not have finished without a supportive and kind supervisor like you. I don't think I could have asked for a better supervisor, mentor and friend. I also really admire your integrity and hard work. Honestly, I'm so happy for your success in the ARC Training Centre in Future Crop Development Fund! You also taught me about priorities in life in how you prioritise your family and other things in your life. I hope that at least a little of your character has rubbed off on me!

Jalaja (Meachery Bhaskaran Jalajakumari), you are the best lab-mum and I'm so grateful to have you by my side for so many years! Thank you for supporting me through the worst and celebrating with me through the best (there were many of both!). Your overwhelming kindness, love and genuine care for me and everyone has played a vital part in this lab feeling like a family and being such a great place. Thank you for caring for me like a daughter and walking right next to me (often literally in the lab). A PhD can be a lonely journey, but because of you, I never felt lonely. I'm so glad we got to share our lives together and hope that we can continue to do so.

To my current lab mates, Leila Blackman, Allan Lohe, Ejiroghene Evivie, Lauren Crean and our honorary member, Ayesha Wellawatta Mudiyanselage. Thanks for making this lab a wonderful place to be! I will miss you guys and I wish you all the best! There have been so many people that have come and gone throughout my time here. I'm especially thankful for my past lab mates, Mohsen Asadi, Naiqi (Nathan) Wang, Herbert Wong, Ke (Luke) Lu, and Laozixian (Zach) Wang. Thank you for being part of my life even if just for a season. This lab environment has

made all the difference in helping me persevere. I really appreciate our time together and I am glad to have known you all. Thank you everyone for being part of the lab family!

I also thank Zhi-Ping Feng whose help was instrumental to me in finishing this thesis. Thank you for teaching me bioinformatics from scratch! Thanks also to Terry Neeman for helping me with statistics so patiently. You were both so encouraging. Thank you to all the staff at the Research School of Biology (RSB) especially to the Plant Services team who were always so friendly and prompt to help even when I made requests last minute. Thanks also to the RSB level 2 kitchenette people. I enjoyed our chats during our unintentionally synchronised lunch and afternoon tea times even though many of you are from the “dark side” (Biomedical Science and Biochemistry) of the building.

Thanks to ANU Women’s Football Club for keeping me sane and mentally and physically healthy. To my team, “you guys, rock!”. Thanks also to Fellowship of Christian University Students (FOCUS), especially to the postgraduate and international student group. There’s too many of you to name but, from the postgraduate group, I especially thank Stephen Driscoll, Chichi Soboya, Brandon Yip, Angus Rae and Wein Lau. I enjoyed our fellowship which was especially important during the COVID-19 lockdowns. From FOCUS international, I especially thank Owen Chadwick and Christine Zheng. I’m so grateful that I’ve been a part of FOCUS for so long. Thanks for keeping me spiritually healthy and reminding me that Jesus comes first in my life.

Thank you, mum and dad, for your love and support through everything especially in the toughest times. If it wasn’t for you I don’t think I would have returned to finish my PhD. You respected my decision either way and, upon deciding to return, you walked beside me every step of the way. Thank you for encouraging me and “waiting it out” with me even though it took me so long to finish! You believed in me even when I didn’t believe in myself. Thank you also to my aunt, Angela (Wah Yee) Cheung for loving me as your niece. Thank you also for believing in me. To Sherman, Queenie, and my beautiful niece, Grace, thank you for being my family and bringing me such joy! (Sherman, I will live my academic life vicariously through you!). Thank you to my friends, Steph Cheung, George Huang, Sheri Kim, Chris Lam, Lily Kim and Jacinta Joe. You guys make me laugh, enrich my life and remind me that there’s life outside of PhD!

I have so much to be thankful for and most of all I thank God. Thank you for giving me this opportunity and, as difficult as it was, You were the one who put all these people around me and gave me the strength to finish.

Abstract

In plants, microRNAs (miRNAs) are short non-coding RNAs of approximately 20-24 nt in length which are involved in post-transcriptional regulation of genes controlling many fundamental biological pathways. They guide the miRNA Induced Silencing Complex (miRISC) to bind to target mRNAs of high complementarity where they are negative regulators of gene expression, acting via transcript cleavage and/or translational repression mechanism(s). Identifying functional miRNA-target interactions (MTIs) is central to understanding miRNA function, and this has led to the development of many miRNA target prediction tools. As high miRNA-target complementarity is required for a MTI in plants, complementarity has been a central factor of these miRNA target prediction tools. However, most of these tools result in long lists of targets, for which there is no experimental evidence supporting the MTI, suggesting the majority of predicted targets are false positives. Furthermore, the degree of complementarity is often used to rank the likelihood of a predicted target as a miRNA target, however, many exceptions have been found. These limitations have impeded our understanding of miRNA biology and the functional scope of miRNA-mediated regulation in plants.

In this thesis, bioinformatic workflow is developed named TRUEE (Targets Ranked Using Experimental Evidence) that ranks MTIs on the extent to which they are subjected to miRNA-mediated cleavage. It sorts predicted targets into high (HE) and low evidence (LE) groupings based on the frequency and strength of miRNA-guided cleavage degradome signals across multiple degradome experiments. From this, each target is assigned a numerical value, termed a *Category Score*, ranking the extent to which it is subjected to miRNA-mediated cleavage. As a proof-of-concept, the 428 Arabidopsis miRNAs annotated in miRBase were processed through the TRUEE pipeline to determine the miRNA “targetome”. The vast majority of high-ranking *Category Score* targets corresponded to highly conserved MTIs, validating the workflow. Very few *Arabidopsis*-specific, *Brassicaceae*-specific, or conserved-passenger miRNAs had HE targets with high *Category Score*s. In total, only several hundred MTIs were found to have *Category Score*s characteristic of currently known physiologically significant MTIs. Although non-exhaustive, clearly the number of functional MTIs is much narrower than what many studies claim. Therefore, using TRUEE to numerically rank targets directly on experimental evidence has given insights into the scope of the functional miRNA targetome of Arabidopsis.

As miRNA-target binding site complementarity is not a definitive indicator of a MTI, this suggests that there are other factors involved in miRNA-mediated regulation. To explore this, TRUEE was applied to conserved miRNAs to determine the identity of HE targets across species and to

investigate potential additional factors involved in miRNA-mediated regulation. Firstly, for each conserved miRNA family, HE targets mostly consisted of one conserved primary target family. If an additional (or secondary) HE target family was identified, it was often functionally related to the primary target family. This suggests that a plant miRNA may preferentially regulate genes that are involved in a functionally similar process. Analyses of the miRNA-target mismatch scores of HE and LE targets further supported the notion that complementarity is not an absolute indicator of a strong MTI. To investigate whether sequences beyond complementarity maybe facilitating MTIs, multiple sequences alignments of conserved target gene homologues were performed. In many instances, these alignments found conserved sequences flanking the miRNA-target binding site. Further bioinformatic analysis found that homologues containing these conserved flanking sequences were enriched in HE targets compared to LE targets, suggesting they are facilitating miRNA-mediated regulation. For a subset of these targets, the conserved flanking sequences were predicted to form conserved RNA secondary structures that preferentially involved base-pairing with the miRNA-bindings sites. This implies many of these conserved miRNA-binding sites are highly structured, counterintuitive to the notion that they should be unstructured and highly accessible for strong miRNA-mediated regulation. Finally, the function of these conserved flanking sequences in the miR160 target, *AUXIN RESPONSE FACTOR 10* (*ARF10*), were functionally tested. The introduction of six synonymous point mutations in the flanking sequences of *ARF10* attenuated its silencing by miR160. Together, these findings suggest that these ancient miRNA-target relationships, have developed regulatory complexities beyond complementarity that define them as strongly regulated target genes of miRNAs.

Table of Contents

Chapter 1

General Introduction

Abbreviations

5’ RACE – 5’ Rapid Amplification of cDNA Ends

AFB – AUXIN SIGNALING F BOX PROTEIN

AGL – AGAMOUS-like

AGO – ARGONAUTE

AMP1 – ALTERED MERISTEM PROGRAM1

AP2 – APETELA2-LIKE

Arabidopsis – *Arabidopsis thaliana*

ARF – AUXIN RESPONSE FACTOR

APS – ATP-SULFURYLASE

Cas9 – CRISPR-associated protein 9

CCS1 – COPPER CHAPERONE FOR SOD1

COX – CYTOCHROME C OXIDASE

CRISPR – clustered regularly interspaced short palindromic repeats

CSD – COPPER/ZINC SUPEROXIDE DISMUTASE

Cu – copper

CUC – CUP SHAPED COTYLEDON

DCL1 – DICER-LIKE1

DRB – DOUBLE-STRANDED RNA-BINDING

ER – endoplasmic reticulum

GRF – GROWTH-REGULATING FACTOR

HAM – HAIRY MERISTEM

HD-ZIPIII – CLASS III HOMEODOMAIN LEUCINE ZIPPER

HYL1 – HYPONASTIC LEAVES

IAR3 – IAA-ALANINE RESISTANT 3

IPS1 – INDUCED BY PHOSPHATE STARVATION1

LAC – LACCASE

LCR – LEAF CURLING RESPONSIVENESS

MBP – membrane-bound polysomes

miRISC – miRNA-induced silencing complex

miRNA – microRNAs

MTIs – miRNA-Target Interactions

N – nitrogen

NF-YA – NUCLEAR TRANSCRIPTION FACTOR Y SUBUNIT ALPHA

NHEJ – nonhomologous end joining

NLA – NITROGEN LIMITATION ADAPTATION

nt – nucleotides

P – phosphate

PHO2 – PHOSPHATE2

PHT5 – PLASMA-MEMBRANE-LOCALIZED PHOSPHATE TRANSPORTER 5

PIWI – P-ELEMENT-INDUCED WHIMPY TESTIS

pri-miRNA – miRNA primary transcript

PTGS – post-transcriptional gene silencing

RISC – RNA-induced silencing complex

RNAi – RNA interference

ROS – reactive oxygen species

S – sulphate

SE – SERRATE

sgRNA – single guide RNA strand

siRNA – small interfering RNA

sRNA – small RNA

STTMs – short tandem target mimics

SULTR2;1 – SULFATE TRANSPORTER2;1

SUO – ‘SHUTTLE’ IN CHINESE

T-plots – target-plots

TAS – TRANS-ACTING SHORT INTERFERING RNA

tasiRNA – trans-acting siRNAs

TCP – TEOSINTE BRANCHED1, CYCLOIDEA, AND PROLIFERATING CELL NUCLEAR ANTIGEN BINDING FACTOR

TGS – transcriptional gene silencing

TIR1 – TRANSPORT INHIBITOR RESPONSE 1

TRUEE – Targets Ranked Using Experimental Evidence

1.1 Gene silencing

RNA interference (RNAi) is a gene regulatory mechanism that is involved in multiple important biological processes in plants, animals and fungi (Dang et al., 2011; Bologna & Voinnet, 2014). Playing a central role in RNAi are small 20-24 nucleotide (nt) regulatory RNAs (sRNA) which associate with an endonuclease, ARGONAUTE (AGO), to form an RNA induced silencing complex (RISC) (Voinnet, 2009; Borges & Martienssen, 2015). The sRNA directs RISC to selectively silence target genes in a sequence-specific manner. In plants, the two major classes of sRNA involved in RNAi are small interfering RNA (siRNA) and microRNA (miRNA) which vary in origin, biogenesis and function (Bologna & Voinnet, 2014). Unlike miRNAs, siRNAs are produced from long double stranded RNAs with near-perfect complementarity. siRNAs regulate via both transcriptional gene silencing (TGS) and post-transcriptional gene silencing (PTGS) depending on the size, origin and biogenesis of the siRNA (Borges & Martienssen, 2015; Wu et al., 2020). Typically, 24 nt-long siRNA regulate gene expression via TGS and maintain genome stability and integrity by directing DNA methylation of transposons and repeats from which it derives (Matzke et al., 2015). 21 and 22 nt-long siRNAs are typically involved in PTGS. Typically, 21 nt siRNAs silence gene expression via target mRNA cleavage and translational repression and 22 nt siRNAs have been found to repress target translation and induce transitive small-RNA amplification (Borges & Martienssen, 2015; Wu et al., 2020). MiRNAs are typically 20-24 nt in length and are produced from imperfectly paired hair-pins. They trigger the PTGS of endogenous genes transcribed from a different locus via transcript cleavage and translational repression (Reinhart et al., 2002; Garcia, 2008). This thesis will predominantly address miRNAs as the main focus.

1.2 miRNAs in plants

In plants, miRNAs are involved in fundamental biological processes such as vegetative and reproductive tissue development, abiotic and biotic stress-responses (Garcia, 2008; Sunkar et al., 2012; Song et al., 2019). In Arabidopsis, miRNAs regulate these processes most commonly by associating with the main effector protein, AGO1, and to a lesser extent AGO2/4/7/10, to form a miRNA induced silencing complex (miRISC) (Reviewed in Song et al., 2019). MiRNAs guide miRISC to bind to target mRNAs of high complementarity where they are down regulated via transcript cleavage and translational repression.

1.2.1 Plant miRNA biogenesis

In plants, miRNAs are typically transcribed by RNA POLYMERASE II. The primary transcript (pri-miRNA) is capped and polyadenylated, then forms an imperfectly paired stem-loop in which the miRNA resides (Reinhart et al., 2002). The miRNA sequence is then excised from the pri-miRNA by a RNaseIII-type endonuclease DICER-LIKE1 (DCL1) to generate a precursor miRNA (pre-miRNA) (Reviewed in Wang et al., 2019). This is achieved in combination with HYPONASTIC LEAVES (HYL1) and SERRATE (SE) which assists the accuracy of pri-miRNA processing, whereby SE acts as a scaffold to recruit HYL1 and RNA substrates to DCL1 for miRNA processing (Kurihara et al., 2006; Machida et al., 2011; Yang et al., 2014). Each of these main effector molecules also function in association with multiple other proteins which positively and negatively regulate pri-miRNA processing (Reviewed in Wang et al., 2019). For example, SMALL1 positively regulates miRNA processing by promoting miRNA gene transcription and promoting DCL1 abundance (Li et al., 2018). The importin β-protein, KARYOPHERIN ENABLING THE TRANSPORT OF THE CYTOPLASMID HYL1, is required for the transport of cytoplasmic HYL1 into the nucleus where miRNA processing occurs (Zhang et al., 2017a). Alternatively, a chromatin remodeling factor, CHROMATIN REMODELLING 2, associates with SE and remodels the RNA substrates to impede processing by DCL1 (Wang et al., 2018). The pre-miRNA is then further processed into a miRNA/miRNA* RNA duplex (henceforth, miRNA guide/passenger strand, respectively) which is then methylated by HUA-ENHANCER 1 where it is protected from degradation (Li et al., 2005; Yu et al., 2005; reviewed in Wang et al., 2019). Typically, the guide strand is then loaded into AGO to form the miRISC and the passenger strand is degraded. The most common model for miRNA loading is that the miRNA duplex is exported from the nucleus and into the cytoplasm by an EXPORTIN5 homologue, HASTY, before loading into AGO1 (Waititu et al., 2020). However, based on the lack of evidence supporting this, the site of miRNA loading remains unclear (Yu et al., 2017; Wang et al., 2019). Contrasting this model, a recent study provided evidence that miRNA loading in AGO1 likely occurs in the nucleus before exportation of the miRISC into the cytoplasm by EXPO1 (Bologna et al., 2018). However, this data does not exclude the possibility that some miRNAs are loaded in the cytoplasm.

1.2.2 miRNA mode of action

Plant miRNA-target interactions (MTIs) require a high degree of complementarity for target transcript cleavage (Reviewed in Yu et al., 2017). MiRNA-target cleavage generally occurs at precise locations on the target binding-site corresponding nt 10-11 of the miRNA (Llave et al., 2002; Jones-Rhoades & Bartel, 2004). Cleavage is achieved by the P-ELEMENT-INDUCED WHIMPY TESTIS (PIWI) domain of AGOs which adopts a RNase H-like fold and possesses endonuclease activity. In *Arabidopsis thaliana* (henceforth, Arabidopsis), AGO1/2/4/7/10 are capable of miRNA target cleavage (Reviewed in Song et al., 2019). After target cleavage, the resulting 5’ and 3’ cleavage product is then each degraded by different pathways ultilising different effector proteins (Souret et al., 2004; Zhang et al., 2017b). In Arabidopsis, the 3’ cleavage product is degraded via a 5’ to 3’ exonuclease, EXORIBONUCLEASE 4 (Souret et al., 2004) and 3’ cleavage product undergoes uridylation followed by degradation by a 3’ to 5’ exonuclease, RISC-INTERACTING CLEARING 3’‐ 5’ EXORIBONUCLEASE (Zhang et al., 2017b). There is emerging evidence which suggests that target cleavage may occur at the rough endoplasmic reticulum (ER). In Arabidopsis, *Zea mays* and *Oryza sativa*, it was found that 21-nt miRNAs were enriched in membrane-bound polysomes (MBP) fractions when compared to total polysomes (Li et al., 2016; Yang et al., 2021). Additionally, transcript fragments corresponding to miRNA-mediated cleavage were generally found to be enriched in these fractions, although in the monocots this varied in different tissues.

Multiple MTIs have been reported to have decreased target protein levels despite little or no changes to mRNA levels indicating translational repression of the target without mRNA decay (Gandikota et al., 2007; Broderson et al., 2008; Li et al., 2014a). Compared to target transcript cleavage, the molecular mechanism underlying translational repression is less clear. Like cleavage, translational repression also requires a high degree of complementarity (Iwakawa & Tomari, 2013). Several effector proteins have been implicated in translational repression but not transcript cleavage. These include ALTERED MERISTEM PROGRAM1 (AMP1), which is an ER membrane protein; VARICOSE (VCS), a cytoplasmic processing (P) body; SUO (‘SHUTTLE’ IN CHINESE), a GW-repeat protein and, KATANIN, a microtubule-severing enzyme (Broderson et al., 2008; Yang et al., 2012; Li et al., 2013a). Mutations to each of these effector proteins resulted in increased target protein levels without affecting mRNA levels. However, how these effector proteins are interrelated and function in translational repression remains unclear and require further studies. It is proposed that the role of VCS and SUO in mRNA decapping and KTN in microtubule dynamics may indicate that these processes are involved in translational repression (Song et al., 2019). AMP1 was also found to be an integral ER membrane protein which may suggest that translational repression occurs at the ER (Li et al., 2013a). It was also found that AMP1 is not required for translational repression for all miRNAs (Fouracre et al., 2020). For instance, *amp1* mutants did not hinder the translational repression of the target genes, *SQUAMOSA PROMOTER-BINDING PROTEIN LIKE* (*SPL*) *9* and *MYB33*, by their miRNAs, miR156 and miR159, respectively. Additionally, AMP1 was found to be an integral membrane protein localised to the rough ER which may suggest that translational repression occurs at this location (Li et al., 2013a). Supporting this is that miRNA transcripts were enriched in MBP fractions in *amp1 lamp1* double mutants (where LAMP1 is an AMP1 paralogue) (Li et al., 2013a). This may suggest that AMP1 prevents miRNA transcript association with the MBP thereby inhibiting translation, however other possibilities exist and require further clarification (Yu et al., 2017). It is clear is that miRNA-mediated regulation operates in a combination of both mechanisms (Gandikota et al., 2007; Broderson et al., 2008; Li et al., 2013a; Li et al., 2014a). However, how this occurs and how much transcript cleavage and translational repression contributes to miRNA-mediated regulation in plants is unknown. A study by Reis et al. (2015) has implicated HYL1 (also known as DOUBLE-STRANDED RNA-BINDING1; DRB1) and DRB2, which are both DCL1 partnering proteins, in determining if a target is translationally repressed or cleaved. DRB2 promotes translational repression by inhibiting HYL1 expression, a promoter of transcript cleavage. However, how DRB2 promotes translational repression over transcript cleavage requires further investigation.

1.2.3 Conservation of plant miRNAs

The miRNA gene regulatory mechanism is highly conserved and has been reported across multiple major plant lineages, including, angiosperms, gymnosperms, lycophytes and non-vascular plants (Floyd & Bowman, 2004; Chávez Montes et al., 2014; You et al., 2017). The conservation of individual miRNA families, however, vary considerably.

Some miRNA families are deeply conserved across hundreds of millions of years. Conserved miRNAs also have highly conserved target families and in any given species, typically both the target and the miRNA belong to gene families consisting of multiple paralogues (Floyd & Bowman, 2004; Axtell and Bartel, 2005; Axtell et al., 2007; You et al., 2017). Furthermore, miRNA abundance correlates with conservation, i.e., the most highly conserved miRNAs also correspond to the most abundant miRNAs composing the majority of total miRNA in a plant (Chávez Montes et al.,2014; You et al., 2017). Reducing the abundance of these conserved miRNAs disrupted their biological function indicating that their abundance is important for miRNA function (Todesco et al., 2010; Yan et al., 2012).

Although the conserved miRNAs constitute the bulk of all expressed miRNAs in a plant, the number of individual conserved miRNA families make up only a minority of all miRNAs in land plants. Rather, the diversity of miRNAs mostly consist of evolutionarily recent or “young”, low abundance, species specific miRNA families with few or only one family member (Rajagopalan et al., 2006; Fahlgren et al., 2007; Chávez Montes et al.,2014). From this observation, it was proposed that young miRNAs emerge frequently which provides a large pool from which new potential MTIs of functional significance can arise (Rajagopalan et al., 2006; Fahlgren et al., 2007; Axtell, 2008). In some instances, a MTI of functional significance arises leading to selective pressure and the retention of the miRNA. One example is the evolutionarily young, *Brassicaceae* specific miR824:*AGAMOUS-like* (*AGL) 16* module which is involved in regulating the number of stomatal complexes, flowering time and heat stress response (Kutter et al., 2007; Hu et al., 2014; Szaker et al., 2019). However, it is proposed that for the majority of these potential miRNAs their presence or absence does not incur a benefit or detriment to plant function. Furthermore, the low abundances of these miRNAs and the difficulty in identifying their targets also suggest their lack of functional significance. Having little or no function, these MTIs are not under strong selective pressure and so undergo neutral genetic drift until the pri-miRNA is no longer recognised by DCL for processing (Rajagopalan et al., 2006; Fahlgren et al., 2007; Axtell, 2008).

1.3 Plant miRNA function

1.3.1 Plant development

As stated above, miRNA-mediated regulation plays crucial and diverse roles in plants. The functional importance of miRNAs in plants was first demonstrated in mutants of core component of the miRNA pathway (*ago1*, *hen1*, *dcl1*, *hyl1, se*) that resulted in pleiotropic mutant phenotypes (Bohmert et al., 1998; Lu & Fedoroff, 2000; Clarke et al., 2002; Park et al., 2002; Reinhart et al., 2002; Han et al., 2004; Vazquez et al., 2004). Similarly, disruption to individual miRNA-target modules can also lead to severe morphological defects (Mallory et al., 2004a; Mallory et al., 2005; Todesco et al., 2010). Correspondingly, many targets of highly conserved miRNAs are regulatory genes that are fundamental for plant function, such as transcription factors and F-box proteins, and control a multitude of downstream genes (reviewed in Jones-Rhoades, 2012).

miRNAs control multiple aspects of plant development. For example, the highly conserved miRNAs, miR156 and miR172, work synergistically to regulate juvenile-to-adult phase transition and flowering (Aukerman & Sakai, 2003; Wu & Poethig, 2006; Wang et al., 2009). They target the transcription factor families, *SPL* and *APETALA2*-like (*AP2*-like), respectively. Overexpression of miR156 prolonged the juvenile phase and delayed flowering in Arabidopsis, *Zea mays* and *Nicotiana tabacum* while, overexpression of the *SPL* family members, *SPL3/4/5*, led to accelerated adult phase transition in Arabidopsis (Wu & Poethig, 2006; Chuck et al., 2007; Zhang et al., 2015). In contrast, overexpression of miR172 and the subsequent downregulation of *AP2*-like target genes in Arabidopsis causes early flowering (Aukerman &

Sakai, 2003). miR156 was found to be most highly expressed during the juvenile phase and decreases before transition into flowering while the opposite is true of miR172 thus indicating a complementary role between these miRNAs. This relationship is also conserved in *Zea mays* and *Oryza sativa* (Chuck et al., 2007; Tanaka et al., 2011). Furthermore, *SPL* expression was also found to be modulated by three miR171 targets from the *GRAS* family, *HAIRY MERISTEM* (*HAM*) *1/2/3* (also known as *LOST MERISTEMS 1/2/3 or SCARECROW-LIKE 6/22/27*) in Arabidopsis (Xue et al., 2014).

miRNAs are also involved in reproductive organ development. miR164 is involved in organ separation and boundary formation via the regulation of *CUP SHAPED COTYLEDON* (*CUC*)*1* and *CUC2* (Laufs et al., 2004). Plants overexpressing miR164 resulted in lowered *CUC1* and *CUC2* levels and displayed separation defects in sepals and stamens and reduced fertility (Laufs et al., 2004; Mallory et al., 2004a). miR159 targets, *MYB33* and *MYB65*, have redundant function in stamen development as *myb33 myb65* double mutants display male sterility (Millar & Gubler, 2005). Similarly, miR159 overexpression also led to anther defects and male sterility (Achard et al., 2004).

In vegetative tissues, a complex regulatory network of miRNAs is also involved in leaf development. One of the major miRNAs is miR165/166 which is involved in the development of the shoot apical meristem via targeting the *CLASS III HOMEODOMAIN-LEUCINE ZIPPER* (*HD-ZIP III*) family genes where the dysregulation to this MTI leads to changes in leaf polarity (Kidner & Martienssen, 2004; Mallory et al., 2004b). Like in floral tissue, miR164 is also involved in the organ boundary formation in vegetative tissues. miR164 overexpressing plants displayed organ separation defects such as, fused cotyledons, fused rosette leaves and fusions of rosette leaves to the stem, and the stem to the pedicle (Laufs et al., 2004; Mallory et al., 2004a). Furthermore, the miR396 target family, *GROWTH REGULATING FACTORs* (*GRFs*), may also interact with *CUC*s to establish organ boundary formation (Lee et al., 2015). In this study, a *cuc* mutant crossed with *grf1/2/3* plants were found to have dramatically more fused cotyledon phenotypes compared to a single *cuc* mutant.

1.3.2 Abiotic stress response

Similarly, miRNAs regulate in response to diverse range of abiotic stress conditions. This includes in multiple nutrients deficiencies. One of the most well studied are miRNAs in copper (Cu) stress response which involve multiple miRNA-target modules. Three widely conserved miRNAs, miR398, miR397 and miR408, are induced by Cu deficiency (reviewed in Pilon, 2016). The targets of these miRNAs are mainly Cu-containing proteins where miR398 targets

COPPER/ZINC SUPEROXIDE DISMUTASE (*CSD*) *1* and *2*, and the Cu chaperone, *COPPER CHAPERONE FOR SOD1* (*CCS1*); miR397 targets *LACCASE* (*LAC*), and miR408 targets *PLANTACYANIN*. In Arabidopsis, these Cu-miRNAs also includes miR857 which also targets *LAC* (Zhao et al., 2015). A study by Shahbaz & Pilon (2019) found that dysregulation of Cu-miRNA-target modules led to reduced plastocyanin levels, a Cu-containing protein involved in photosynthesis, and *CYTOCHROME C OXIDASE* (*COX*), which is involved in the respiratory electron transport chain. Therefore, this suggests a key role for the down-regulation of these Cu-containing proteins is to direct Cu for the most important Cu-containing proteins when Cu is limited. MiRNAs also play a role in response to sulphate (S) deficiency. Under these conditions, the highly conserved miR395 is induced and targets *ATP-SULFURYLASE* (*APS*) *1*, *3* and *4* and *SULFATE TRANSPORTER2;1* (*SULTR2;1*) which are both involved in S metabolism and transport (Jones-Rhoades & Bartel, 2004; Liang et al., 2010). Further examples include the involvement of miR399:*PHOSPHATE2* (*PHO2*) in phosphate (P) deficiency (Chiou et al., 2006); miR169:*NUCLEAR TRANSCRIPTION FACTOR Y SUBUNIT ALPHA* (*NF-YA*) in response to nitrogen (N) deficiency (Zhao et al., 2011); and miR827: *NITROGEN LIMITATION ADAPTATION* (*NLA*) in Arabidopsis (or *PLASMA-MEMBRANE-LOCALIZED PHOSPHATE TRANSPORTER 5* (*PHT5*) in other angiosperms) which is also in response to phosphate (P) deficiency (Lin et al., 2018).

Apart from nutrient deficiency, miRNAs are also implicated in other abiotic stresses. For example, miR398 is also reported to participate in multiple stresses including UV-light, heavy metal, methyl viologen-induced oxidation (Sunkar et al., 2006). These conditions cause oxidative stress generating reactive oxygen species (ROS). Here, miR398 is repressed for the upregulation of CSD1 and CSD2 which scavenge for ROS leading to superoxide detoxification (Sunkar et al., 2006). Demonstrating this, overexpressing a miR398 resistant *CSD2* led to greater tolerance under high-light, high Cu stress and methyl viologen induced oxidative stress (Sunkar et al., 2006). In response to drought stress, down-regulation of miR167 derepresses its target, *IAA-ALANINE RESISTANT 3* (*IAR3*), which results in lateral root development and drought resistance in Arabidopsis (Kinoshita et al., 2012). Additionally, in a study by Li et al. (2008), miR169 was also shown to be down-regulated under drought stress leading to increased *NF-YA5* levels. A *nf-ya5* knockout plant and *NF-YA5* over-expressing plant were also generated in this study finding these plants to have poorer and greater drought resistance, respectively. Under UV-irradiation, miR396 is induced which silences *GRF1*, *GRF2* and *GRF3* under UV-irradiation and leads to the inhibition of cell proliferation (Casadevall et al., 2013). In response to cold stress, several miRNAs are induced in different species leading to target regulation and enhanced cold tolerance. This includes miR393 regulation of *TEOSINTE BRANCHED1, CYCLOIDEA, PROLIFERATING CELL NUCLEAR ANTIGEN BINDING FACTOR* (*TCP*) in sugarcane and rice (Thiebaut et al., 2011; Yang et al., 2013); miR393 regulation of *TRANSPORT INHIBITOR RESPONSE 1* (*TIR1*) and *AUXIN SIGNALING F BOX PROTEIN* (*AFB*) in switchgrass and (Liu et al., 2017a); and miR394 regulation of *LEAF CURLING RESPONSIVENESS* (*LCR*) in Arabidopsis (Song et al., 2016).

1.3.3 Biotic stress response

Plants deficient in miRNA activity (*ago1*) also display poorer pathogen resistance thus implicating a role of miRNAs in biotic stress response (Morel et al., 2002). Across a variety of species, different miRNAs were found to be up or down regulated in response to different pathogens (Huang et al., 2016). One of the first demonstrations of this was in Arabidopsis where leaves challenged with flagellin-derived peptides were found to induce miR393 which led to the downregulation of its targets, which are from a family of F-box auxin receptors, and increased resistance to *Pseudomonas syringae* (Navarro et al., 2006). This was later found to be due to the targeting of the *TCP* family member, *AFB1*, which alleviated the repression of *SALICYLIC ACID* by *AUXIN RESPONSE FACTORS* (*ARF*) leading to resistance to biotropic and hemibiotropic pathogens (Robert-Seilaniantz et al., 2011). Similarly, miR393 was also induced in other plant species (*Nicotiana tabacum*, *Glycine max* and *Zea mays*) when challenged with biotropic and hemibiotropic pathogens (Reviewed in Šečić et al., 2021). However, miR393 was downregulated in *Solanum melongena* when challenged with a necrotropic pathogen, *Verticillium dahliae,* which activates a pathway antagonistic to *SALICYLIC ACID* signalling. Similarly, miR160 and miR167, which directly target *ARF* family genes were either upregulated or downregulated in response to biotropic and necrotropic pathogen across different species (Reviewed in Šečić et al., 2021).

1.3.4 miRNA homeostasis

Plant miRNAs also function in maintaining miRNA homeostasis by targeting key effector proteins involved in miRNA biogenesis and action resulting in negative feedback regulation. *AGO1* fragments corresponding to the cleavage products from a predicted miR168-binding-site in *AGO1* were first identified by Vaucheret et al. (2004a). In a subsequent study, they found that plants over-expressing a miR168-resistant *AGO1* gene displayed a loss-of-function *ago1* phenotype and displayed some pleiotropic defects reminiscent of *dcl1*, *hyl1* and *hen1* mutants (Vaucheret et al., 2004). These mutant phenotypes could then be rescued by introducing an artificial miRNA that had high complementarity to the miR168-resistant *AGO1*, demonstrating a negative feedback loop that is required to maintain miRNA homeostasis. Another major effector protein in miRNA biogenesis, *DCL1*, also undergoes miRNA-mediated regulation. For *DCL1*, it was found that *DCL1* levels were elevated in *dcl1* and *hen1* mutant plants and, in contrast, were found at relatively low levels in WT plants with a functional DCL1 (Xie et al., 2003). Furthermore, *DCL1* was also found to have a miR162 binding-site for which miR162-guided cleavage products were found. In Arabidopsis, miR838 was also found to derive from the 14th intron *DCL1* which resulted truncated *DCL1* fragments (Xie et al., 2003; Rajagopalan et al., 2006). Together, this is consistent with the model that *DCL1* undergoes negative feedback regulation via miRNA-mediated regulation.

1.4 trans-acting siRNAs also play crucial functions in plants

Corresponding to another class of siRNA, trans-acting siRNAs (tasiRNA) are 21-22 nt siRNA that appear to only be found in plants. They have overlapping features of both siRNA and miRNA biogenesis and function. Reminiscent of siRNA, tasiRNA are generated from a single stranded precursor tasiRNA transcript. This precursor is then targeted and cleaved by a 21 or 22 nt miRNA and converted into dsRNA by RNA DEPENDENT RNA POLYMERASE 6 (Peragine et al., 2004; Allen et al., 2005). The ds-tasiRNA precursor is then processed into 21-nt phased products by DCL4 in register of the miRNA cleavage site (Gasciolli et al., 2005). Similar to miRNAs, tasiRNAs act in *trans* and target endogenous gene families transcribed at different loci. Although four tasiRNA loci are known in Arabidopsis *TRANS-ACTING SHORT INTERFERING RNA* (*TAS) 1-4*, only *TAS3* will be studied in this thesis (Allen et al., 2005; Rajagopalan et al., 2006). *TAS3* has two miR390 binding-sites on the 5' and 3' end of the transcript. A miR390-AGO7 complex executes cleavage of the 3' binding-site to set the register for the correct tasiRNA products that then targets *ARF2*/*ARF3*/ *ARF4* which are involved in vegetative phase change, leaf patterning and lateral root growth in Arabidopsis (Peragine et al., 2004; Williams et al., 2005; Adenot et al., 2006; Fahlgren et al., 2006; Garcia et al., 2006; Hunter et al., 2006; Montgomery et al., 2008; Marin et al., 2010). Similar to many miRNAs that are crucial for plant development, the miR390-*TAS3-ARF* module is also highly conserved, spanning to the nonvascular land plant *Marchantia polymorpha* (Xia et al., 2017).

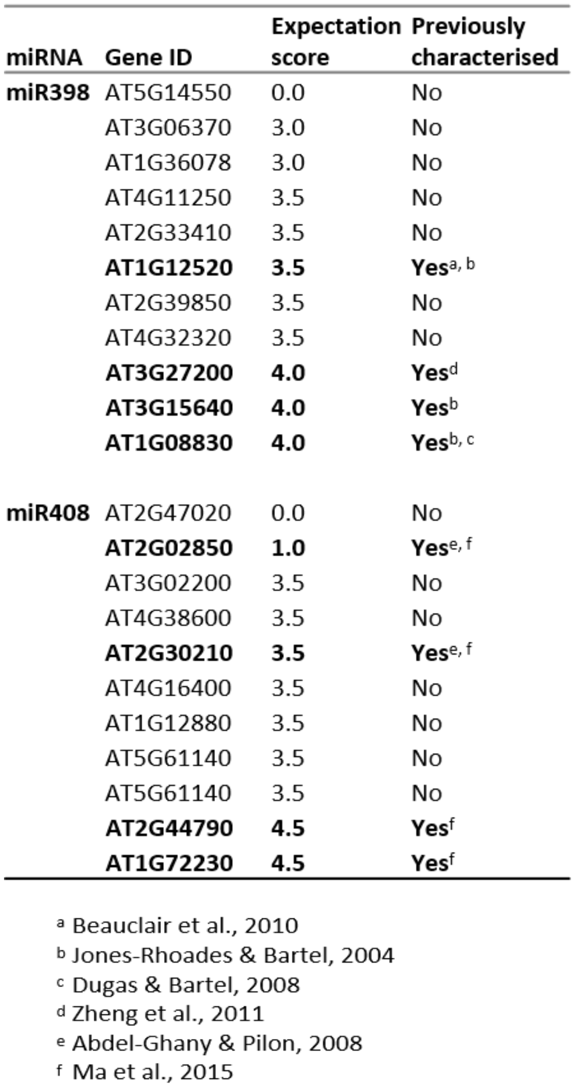
1.5 Experimental approaches to identify miRNA target genes in plants

1.5.1 Bioinformatic prediction

The identification of a miRNA's targets is integral in determining its function. As previously mentioned, a high degree of complementarity is required for miRNA-mediated regulation between the miRNA and its target binding-site in plants (Schwab et al., 2005). As such, many bioinformatic programs have been developed which predicts targets based on miRNA-target complementarity, and have been highly successful in identifying miRNA targets that subsequently were experimentally validated (Jones-Rhoades & Bartel, 2004). These studies observed that the number and positions of mismatches at the miRNA-target binding-sites impacted miRNA-mediated regulation. Consequently, this led to the development of mismatch scoring schemas to bioinformatically predict genes with the highest confidence as miRNA targets. Initially, Rhoades et al. (2002) considered genes that were complementary with ≤ 3 mismatches to the miRNA to be likely targets. However, this was unable to identify many other miRNA targets that were previously experimentally validated. Rhoades and Bartel (2004) added a mismatch penalty score that considered the type of mismatch (G:U pair = 0.5, non-G:U pair = 1), and bulged nucleotides/gaps in either the miRNA or target binding-site (= 2). This scoring schema was further refined by considering the positions of mismatches. Fewer mismatches were found to be tolerated at the 5' end of the miRNA (nt positions 2-13) and therefore led to a heavier mismatch score weighting (Mallory et al., 2004a; Allen et al., 2005; Schwab et al., 2005). Additionally, no mismatches were tolerated between the 10-11 nt target cleavage site of the miRNA (Mallory et al., 2004a; Schwab et al., 2005). Subsequently, complementarity-based scoring schemas became a central component in miRNA target prediction tools, such as the widely used psRNATarget and TAPIR, which often ranked the confidence of a gene as a miRNA targets via the degree of complementarity (Bonnet et al., 2010; Dai & Zhao, 2011; Sun et al., 2011; Dai et al., 2018).

However, such bioinformatic programs resulted in a long list of predicted targets and ranking the likelihood of a gene as a target by complementarity was not able to consistently predict an experimentally validated target resulting in many false positives. For example, of a group of eight *GAMYB-like* genes that all contained conserved miR159-binding-sites, only two of these genes (*MYB33* and *MYB65*) were strongly silenced. This was despite *MYB33* and *MYB65* having a higher Expectation Score (the mismatch score metric from the target prediction program, psRNATarget) than other *GAMYB-like* genes (e.g. *MYB104* and *MYB101*) which were poorly regulated (Zheng et al., 2017). Furthermore, miRNAs such as miR398 and miR408 have canonical targets with high mismatch scores, while some predicted targets with lower scores, based on the absence of literature, have not been experimentally validated (Table S1) (Table

1.1). To improve the number of functional targets identified while reducing false positives, some bioinformatic studies incorporate conservation as a filter and have been successful in identifying miRNA targets that have been previously overlooked. Moreover, this approach was able to reduce false positives (Chorostecki et al., 2012; Chorostecki et al., 2014; Ma et al., 2018). This is based on the principle that biologically advantageous MTIs would be selected for during evolution and conserved.



miRNA	Gene ID	Expectation score	Previously characterised
miR398	AT5G14550	0.0	No
	AT3G06370	3.0	No
	AT1G36078	3.0	No
	AT4G11250	3.5	No
	AT2G33410	3.5	No
	AT1G12520	**3.5**	**Yes**[a, b]
	AT2G39850	3.5	No
	AT4G32320	3.5	No
	AT3G27200	**4.0**	**Yes**[d]
	AT3G15640	**4.0**	**Yes**[b]
	AT1G08830	**4.0**	**Yes**[b, c]
miR408	AT2G47020	0.0	No
	AT2G02850	**1.0**	**Yes**[e, f]
	AT3G02200	3.5	No
	AT4G38600	3.5	No
	AT2G30210	**3.5**	**Yes**[e, f]
	AT4G16400	3.5	No
	AT1G12880	3.5	No
	AT5G61140	3.5	No
	AT5G61140	3.5	No
	AT2G44790	**4.5**	**Yes**[f]
	AT1G72230	**4.5**	**Yes**[f]

[a] Beauclair et al., 2010
[b] Jones-Rhoades & Bartel, 2004
[c] Dugas & Bartel, 2008
[d] Zheng et al., 2011
[e] Abdel-Ghany & Pilon, 2008
[f] Ma et al., 2015

Table 1.1. List of psRNATarget predicted targets for miR398 and miR408. Expectation Scores of miR398 and miR408 predicted targets from psRNATarget in Arabidopsis (Dai et al., 2018). Expectation Score is the mismatch score by psRNATarget where a lower score denotes higher complementarity. Bolded targets indicate canonical targets which have been previously experimentally validated. Unbolded targets are predicted targets for which there is yet to be experimental evidence validating it as a target. These do not represent comprehensive lists of predicted targets from psRNAtarget.

1.5.2 Degradome analysis

Experimental validation is required to support bioinformatic prediction. A method developed by Llave et al. (2002) paved the way in identifying cleaved miRNA targets experimentally. This method used a modified RNA Ligase mediated 5’ Rapid Amplification of cDNA Ends (5’ RACE) protocol. As miRISC typically cleaves targets between the 10-11 nt position relative to the 5’ end of miRNA, this leaves a 5’-monophosphate which can be ligated with a 5’-RNA adaptor. Sequencing can then be used to determine the location of the cleavage site. This method has been very successful in identifying cleaved miRNA targets and remains one of the most widely used methods to date. However, 5’RACE can only analyse one potential target at a time and requires prior knowledge of its sequence.

A transcriptome-wide extension of this is degradome sequencing (Addo-Quaye et al., 2008; German et al., 2008; German et al., 2009), which also utilizes 5’-RNA adaptor ligation onto uncapped transcripts, followed by sequencing to determine the precise miRNA cleavage site (Figure 1.1; Figure 1.2). Degradome sequencing captures all polyadenylated uncapped transcripts thus providing a high-throughput global snapshot of all degraded transcript. As degradome sequences captures transcripts in parallel, results also reflect relative abundances of these transcripts and therefore the extent to which a target is cleaved. In contrast, 5’-RACE individually amplifies targets in isolation by PCR and, therefore, may be detecting inefficient basal cleavage activity of little functional significance. For example, there are around twenty predicted miR159 targets of which ten have been validated by 5’-RACE. However, only two of these, *MYB33* and *MYB65*, have been demonstrated to be functionally significant via genetic analysis *in planta* (Allen et al., 2007, 2010). Consistent with this, out of all the *MYB* genes, degradome sequencing also only consistently detected degradome reads reflective of strong miR159-mediated cleavage for *MYB33* and *MYB65* (Addo-Quaye et al., 2008; German et al., 2008). The addition of degradome sequencing data in identifying miRNA targets has refined the long list of bioinformatically predicted targets by providing experimental evidence of target transcript cleavage and have been thenceforth incorporated into multiple miRNA target prediction programs (Addo-Quaye et al., 2009; Zheng et al., 2012; Folkes et al., 2012; Ma et al., 2018).

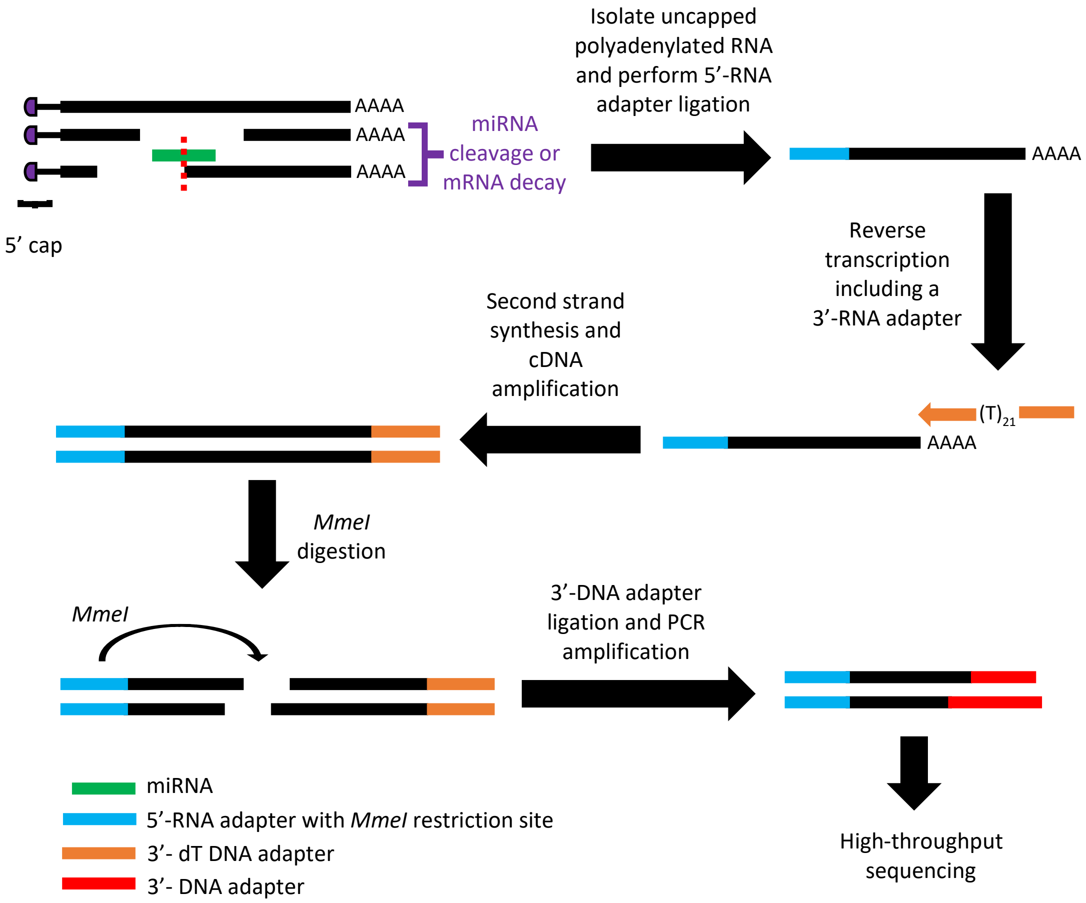


Figure 1.1. Degradome library construction workflow. The steps for degradome library preparation are as follows: 1) isolate polyadenylated RNA; 2) ligate the 5’-RNA adapter to the 5’-monophosphate of 3’ truncated RNA (as a product of miRNA-mediated cleavage or mRNA decay). The adaptor has a *MmeI* restriction site; 3) perform a reverse transcription for cDNA first strand synthesis using an oligo(dT) primer with a 3’ adaptor sequence; 4) conduct a PCR for second strand synthesis and cDNA amplification; 5) perform a *MmeI* digestion to generate equal 20 bp sized sequences; 6) ligate the 3’-DNA adapter with degenerate sites to *MmeI* digestion products; 7) Use PCR amplification then gel purification and submit the degradome library for high-throughput sequencing.

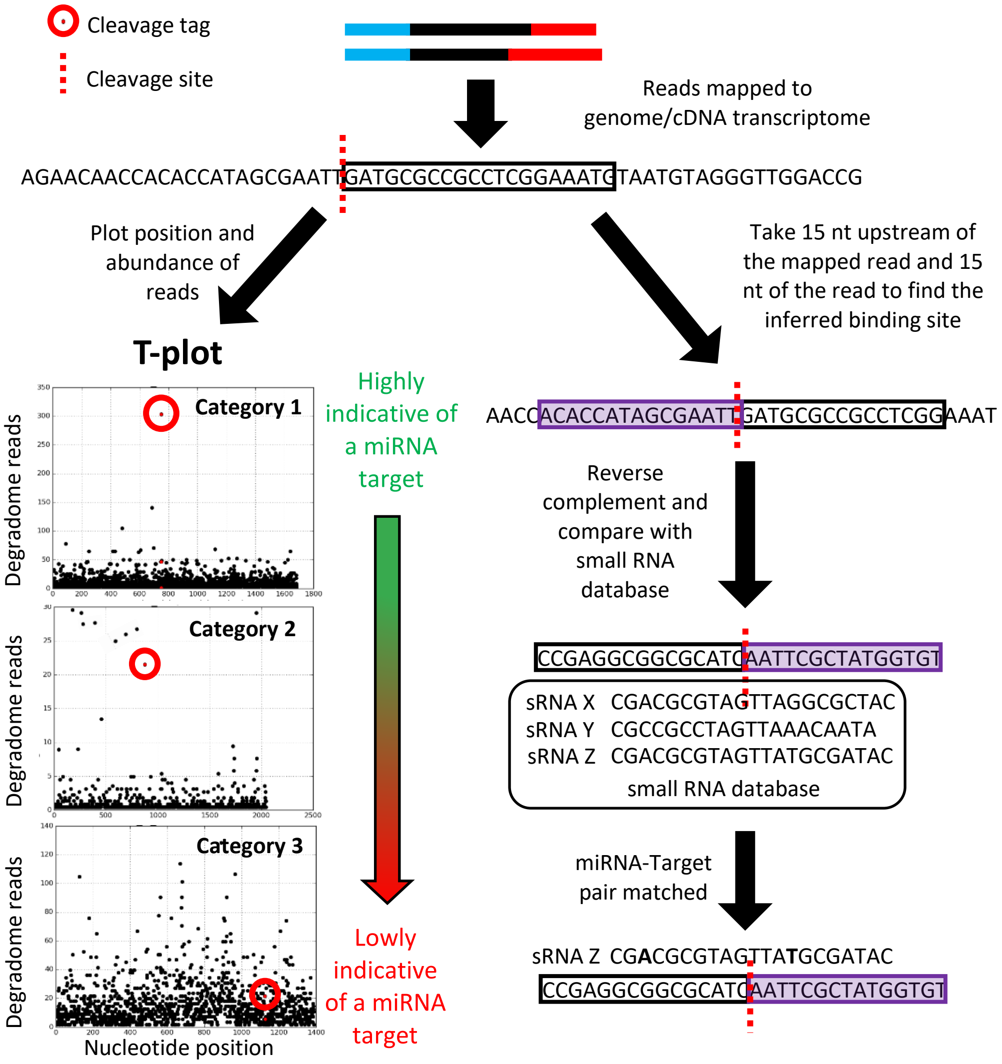


Figure 1.2. Analyses of degradome sequencing results to identify miRNA targets. The steps to identify miRNA targets using bioinformatics analysis is as follows (Addo-Quaye et al., 2008; German et al., 2009): 1) remove the adapter sequences and map reads to the genome/cDNA transcriptome. 2) Plot the position and abundance of reads along gene transcripts to generate target-plots (T-plots). T-plots compare the relative abundance of reads mapping exactly to potential cleavage site (cleavage tag) (corresponding to nts 10-11 relative to the miRNA) from all other reads on the transcript to distinguish it from background noise. The higher the abundance of the cleavage tag from other reads, the more likely it is to indicate miRNA-mediated cleavage. To filter for likely cleavage sites, T-plots are categorised into Categories 1-4 which respectively indicates most to least likely cleaved by a miRNA (only Category 1-3 shown). The above T-plots are adapted from WPMIAS (Fei et al., 2020). 3) To identify the miRNA corresponding to the cleavage tag, infer the whole target binding site by adding the 15 nts upstream of the mapped read to the whole or first 15 nts of the read (depending on method) to create a 30-36 nt sequence (t-signature). Reverse complement the t-signature and match to a database of miRNAs allowing for several mismatches.

1.5.3 Genetic and transgenic studies

MiRNA targets have also been identified through *in planta* functional genetic studies. Generally, these methods provide stronger lines of evidence to validate MTIs of functional importance, but typically are low throughput and more difficult to perform. Overexpression of the miRNA is often used followed by phenotyping and transcript analysis to study the effects of miRNA overexpression (Mallory et al., 2005; Schwab et al., 2005; Debernardi et al., 2014). However, overexpression of the miRNA at artificially high levels (even using an endogenous promoter) may not be reflective of endogenous cleavage as stochiometric ratios of the miRNA-target pair will influence the regulatory outcome of the MTI (Li et al., 2014a).

Another validation method is the overexpression of a miRNA resistant target (Mallory et al., 2005; Wu et al., 2006; Kutter et al., 2007; Liang et al., 2014). These are miRNA targets which have mutations introduced to the miRNA binding-site so that it is no longer miRNA-regulated, and thus resulting in a miRNA loss-of-function-like phenotype. As this also requires overexpression, this may lead to ectopic expression and artificially high levels of the target mRNA leading to exaggerated phenotypic defects that do not reflect miRNA regulation (Li et al., 2014a). A miRNA loss-of-function method more reflective of endogenous conditions is using T-DNA insertional mutants. However, mutants are difficult to generate due to the small size of miRNA genes and that many miRNAs belong to families of multiple functionally redundant homologues.

Overexpression of miRNA decoys can overcome this redundancy by simultaneously silencing whole families of miRNA homologues (Todesco et al., 2010; Yan et al., 2012; Reichel et al., 2015). MiRNA decoys are typically non-coding RNA designed with high complementarity binding-sites to sequester miRNA activity by competing with targets for miRNA binding. The identification of an endogenous miRNA decoy in plants, *INDUCED BY PHOSPHATE STARVATION1* (*IPS1*), which targets miR390, then led to the design of multiple artificial decoys such as, target *MIMIC*s, *SPONGE*s and short tandem target mimics (STTMs) (Franco-Zorrilla et al., 2007; Todesco et al., 2010; Ivashuta et al., 2011; Yan et al., 2012; Reichel et al., 2015). One of the most widely used miRNA decoy is STTMs which has been successfully employed to study miRNA function across a variety of plant species including agronomically important crops such as *Zea mays*, *Oryza sativa* and *Solanum lycopersicum* (Yan et al., 2012; Zhang et al., 2017c; Peng et al., 2018). They have been used to study both highly conserved and lineage specific miRNAs where many of the latter are poorly studied (Peng et al., 2018). However, decoy efficacies vary greatly when targeting different miRNAs and no one approach can robustly ensure strong sequestration across all miRNAs (Reichel et al., 2015), with factors such as RNA secondary structure of the decoys likely effecting their efficacies (Wong et al., 2018).

Furthermore, as miRNA decoys target the mature miRNA, they are unable to distinguish between miRNAs with similar sequences which may lead to misattributing target genes to a miRNA. For example, *MIMIC*s designed to sequester miR159 were found to also sequester the closely related miR319 and *vice versa* (Reichel & Millar, 2015).

More recently, the clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated protein 9 (Cas9) system, has become increasingly used to study miRNA function across various plant species by generating miRNA knockout mutants (Miao et al., 2019; Xu et al., 2020; Hong et al., 2021; Lian et al., 2021; reviewed in Deng et al., 2022). CRISPR/Cas9 is a precise gene editing mechanism which requires the Cas9 endonuclease and a sequence-specific single guide RNA strand (sgRNA) which directs it to a target locus (Chen et al., 2019). This generates a double-stranded DNA break which then undergoes DNA repair mechanisms such as nonhomologous end joining (NHEJ). NHEJ is error-prone and can introduce insertions or deletions (mostly of 1-3 nts in size), thereby, disrupting gene function. The specificity of CRISPR/Cas9 allows the targeting of individual miRNA genes with similar sequences (Miao et al., 2019; Lian et al., 2021). Five individual miR172 homologues were disrupted separately which enabled the elucidation of the distinct and redundant function of these miRNAs in Arabidopsis (Lian et al., 2021). In addition, the multiplex CRISPR/Cas9 system can simultaneously edit multiple genes thereby overcoming the problem of functional redundancy from multiple miRNA homologues (Miao et al., 2019; Hong et al., 2021). As CRISPR/Cas9 requires an expression vector containing the sgRNA and Cas9 protein cassette, this can be achieved by designing a vector with multiple sgRNAs (Miao et al., 2019; Hong et al., 2021).

1.6 Factors beyond binding-site complementarity

Although a high degree of miRNA-target complementarity is a prerequisite for a strong MTI outcome in plants (Schwab et al., 2005), it is clear that high complementarity does not guarantee one. From almost two decades of study, only a select subset of bioinformatically predicted targets with high complementarity have been experimentally shown to be *bona fide* miRNA targets. From these long lists of predicted targets (Dai et al., 2018), there are validated targets which possess lower complementarities, while there are predicted targets with higher complementarities for which no validation exists (Table 1.1) (Jones-Rhoades & Bartel, 2004; Brousse et al., 2014; Zheng et al., 2017). Using a transient assay system, Liu et al. (2014) found that binding-sites engineered with perfect complementarity were not the most strongly silenced. This was also evident from miRNA-based technologies where designs using a high degree of complementarity did not ensure strong miRNA-mediated regulation. For example, artificial miRNAs (amiRNA) designed with analogous complementarities silenced their targets at varying efficacies (Deveson et al., 2013; Li et al., 2013b). This was also found for miRNA decoys. Different decoys, which varied in size and sequences, designed with identical binding-sites were found to inhibit miRNAs at varying efficacies suggesting the sequence context of the binding-site influence the strength and outcome of the MTI (Reichel et al., 2015; Wong et al., 2018). Altogether, this evidence indicates that factors beyond miRNA-binding-site complementarity are involved in miRNA-mediated regulation.

In animal systems, target site accessibility, RNA-binding proteins and RNA secondary structures near miRNA binding-sites have been demonstrated to be involved in miRNA-mediated regulation (Kertesz et al., 2007; Kedde et al., 2010). Kertesz et al. (2007) demonstrated a role of target site accessibility on miRNA-mediated regulation by introducing mutations in the nucleotides adjacent to the binding-site of known targets to render the binding-site into a highly paired stem-loop. Results found that closed conformations hindered the MTIs at a level comparable to mutations within the miRNA binding-site. Considering the miRNA is bound to an AGO protein, a tight conformation may sterically hinder access to the miRNA binding-site (Figure 1.3).

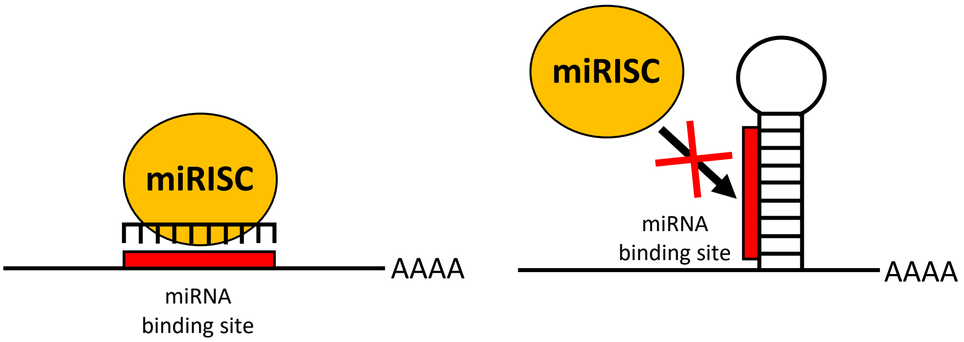


Figure 1.3. A proposed effect of target site accessibility on miRNA target recognition. If the miRNA binding site is in an open RNA secondary structure it is highly accessible to miRISC. Whereas miRISC binding is prevented if it is in a highly structured RNA sequence.

Studies in plants have also implicated a role of target site accessibility and RNA secondary structures in plant miRNA-mediated regulation. A bioinformatics study analysing the genome of four different plant species found an AU rich codon bias in the 96 nt sequences flanking upstream and downstream of the miRNA binding-sites (Gu et al., 2012). This suggests these flanking sequences are under selective pressure to have less RNA secondary structure surrounding these miRNA binding-sites. However, this analysis was performed on targets predicted using psRNAtarget, and so it is unclear of how strong these target genes are regulated by miRNAs. However, an *in vitro* analysis of the RNA secondary structure of the *Arabidopsis thaliana* transcriptome found that the 21 nt miRNA binding-sites of psRNATarget predicted targets was less structured compared to the 50 nt directly flanking upstream and downstream of the binding-site suggesting a greater accessibility of this region (Li et al., 2012). Both these studies suggest that weak RNA secondary structure is a feature of miRNA-binding-sites, potentially making these regions highly accessible to miRISC complexes. However, contrary to these studies, functional studies *in planta* found AU content flanking miRNA-binding-sites did not correlate with stronger miRNA-mediated gene silencing (Deveson et al., 2013; Zheng et al., 2017).

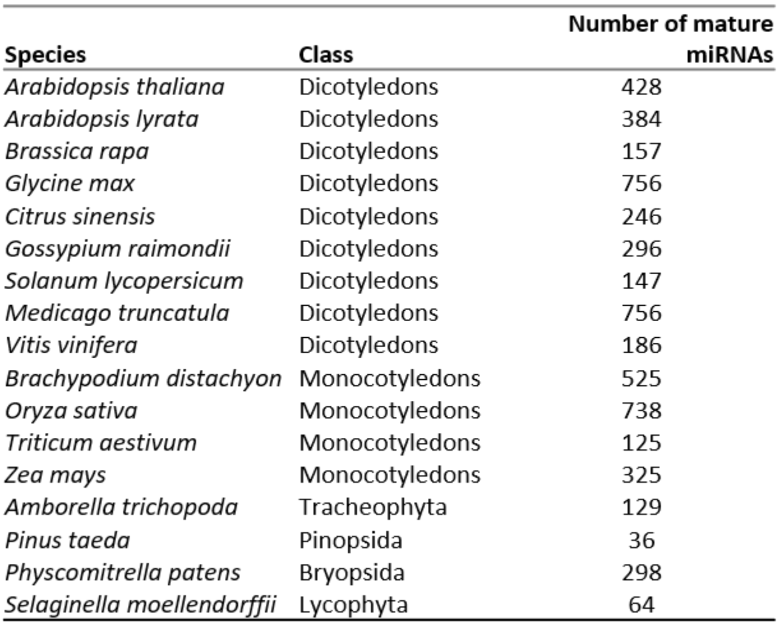
In contrast to the notion that miRNA binding-site need to be devoid of strong RNA secondary structures, an *in vivo* study found the miRNA binding-sites to be more structured and not accessible to miRISC prior to target cleavage (Yang et al., 2020). Rather, the unfolding of this RNA secondary structure acts as a rate-limiting factor of miRISC cleavage efficiency. Here, only the 2 nts immediately downstream of the binding-site were required to be single-stranded for efficient cleavage, although they found that this does not affect miRISC ability to bind to the binding-site. As such, the reason for these unstructured nts differs from the traditional notion that RNA secondary structures dictate spatial accessibility of the binding-site by miRISC.

Supporting the notion of highly structured miRNA binding-sites was the discovery of a highly conserved RNA secondary structures directly upstream of the miR159 binding-site in *GAMYB* genes (Zheng et al., 2017). The nucleotides corresponding to the stems of these RNA secondary structures were found to be conserved across diverse higher plant species, implying the RNA secondary structure is under selective pressure. A structure/function analysis was performed on these putative RNA secondary structures, and mutations that disrupted these RNA secondary structures attenuated miR159-regulation, whereas further compensatory mutations were made to recreate the structures, restore strong miR159-regulation, demonstrating a role for these RNA secondary structures in miRNA-mediated regulation (Zheng et al., 2017).

Although evidently involved in miRNA-mediated regulation in plants, the prevalence and impact of factors beyond binding-site complementarity on miRNA-mediated regulation remains uncertain. As relatively few examples have been described to date, some of which are opposing notions, they cannot be considered general features of MTIs with a high degree of certainty. As such, this remains a challenge in the development of more accurate bioinformatic prediction of miRNA targets with parameters beyond miRNA-target complementarity.

1.7 The functional scope of miRNA-mediated regulation in plants remains contentious

Currently, there generally exists two opposing notions on the functional scope of miRNA-mediated regulation in plants. Many studies point to a complex “miRNome” (the entirety of plant miRNAs), which in turn implies there are potentially thousands of MTIs that confer a plethora of functions (Lindow & Krogh, 2005; Lindow et al., 2007; Meng et al., 2011; Bülow et al., 2012; Fei et al., 2020). This notion is supported by studies suggesting that most miRNAs and MTIs are lineage-specific even between closely related species, suggesting miRNAs are evolutionarily fluid and able to generate many diverse MTIs over a short period of evolutionary time (Smith et al., 2015; Cui et al., 2017). Adding to the complexity of the miRNome, are isomiRs (sequence variants of miRNAs due to altered DCL processing or post-processing modifications), and miRNA passenger strands. Initially, due to their low abundance and the preferential degradation, the miRNA passenger strand was thought to be only a by-product of miRNA biogenesis. However, multiple studies have reported that the miRNA passenger strands have functional roles, and therefore adding another layer of complexity to the miRNome (Reviewed in Liu et al., 2017b). On top of this complexity, advances to sequencing technology over the past decade has led to the identification and annotation of a multitude of low abundance young miRNAs. These have then been uploaded onto miRbase, the largest and most widely used miRNA database (Kozomara et al., 2019). In the latest release (v22), 1000s of different miRNA sequences have been reported across many diverse plant species (Kozomara et al., 2019) (Table 1.2).



Species	Class	Number of mature miRNAs
Arabidopsis thaliana	Dicotyledons	428
Arabidopsis lyrata	Dicotyledons	384
Brassica rapa	Dicotyledons	157
Glycine max	Dicotyledons	756
Citrus sinensis	Dicotyledons	246
Gossypium raimondii	Dicotyledons	296
Solanum lycopersicum	Dicotyledons	147
Medicago truncatula	Dicotyledons	756
Vitis vinifera	Dicotyledons	186
Brachypodium distachyon	Monocotyledons	525
Oryza sativa	Monocotyledons	738
Triticum aestivum	Monocotyledons	125
Zea mays	Monocotyledons	325
Amborella trichopoda	Tracheophyta	129
Pinus taeda	Pinopsida	36
Physcomitrella patens	Bryopsida	298
Selaginella moellendorffii	Lycophyta	64

Table 1.2. Total mature miRNA entries across diverse plant species on miRBase v22. The number of mature miRNAs annotated on miRBase v22 per plants species (Kozomara et al., 2019).

Opposing this hypothesis, other studies have proposed a much narrower functional scope of miRNA-mediated regulation (Meng et al., 2012; Li et al., 2014b; Taylor et al., 2014; Taylor et al., 2017; Axtell & Meyers, 2018). An observation is that the bulk of these lineage-specific miRNAs are poorly conserved, weakly expressed and appear to have no clear target (Rajagopalan et al., 2006; Fahlgren et al., 2007). As such, it was proposed that the majority of young miRNAs are of little biological function and are undergoing neutral drift (Axtell, 2008; Cuperus et al., 2011). Many studies have also questioned the validity and quality of the user-submitted miRNA entries on miRBase, suggesting that a greater majority of these are false positives (Meng et al., 2012; Taylor et al., 2017; Axtell & Meyers, 2018; Kozomara et al., 2019). This is in part due to the lack of adherence to guidelines on correct miRNA annotation (Taylor et al., 2014; Axtell & Meyers, 2018).

1.8 Objectives of thesis

miRNA are master regulators of gene expression and have diverse role in many important plant processes. Key to understanding their function is the identification of their target genes, however, this remains challenging and is limited with current bioinformatic approaches.

Chapter 2 develops an improved bioinformatics pipeline for identifying functionally relevant miRNA targets; Targets Ranked Using Experimental Evidence (TRUEE). The novelty of TRUEE is it uses degradome sequencing data to score potential miRNA targets, rather than miRNA-target binding-site complementarity. From TRUEE analysis, targets are then ranked based on the strength and frequency of their degradome signatures across multiple experiments. This identified which targets are most likely to be subjected to functionally significant miRNA-mediated regulation. This TRUEE pipeline was then applied to analyse all Arabidopsis miRNA reported on miRBase to generate an accurate estimate of the total number of active miRNAs and their complete set of targets (miRNA targetome) in plant for the first time.

Chapter 3 then utilises TRUEE to investigate the identity of target genes of highly conserved miRNA across diverse plant species. These targets are then investigated to determine the extent to which complementarity can be used as an indicator of a strong MTI and whether sequences flanking the miRNA-binding-sites are conserved and correlated with targets that are strongly miRNA-regulated. Finally, the flanking sequences in one highly conserved Arabidopsis target is functionally tested *in planta* to determine its role in miRNA-mediated regulation.

References

Abdel-Ghany, S. E., & Pilon, M. (2008). MicroRNA-mediated systemic down-regulation of copper protein expression in response to low copper availability in Arabidopsis. *Journal of Biological Chemistry*. https://doi.org/10.1074/jbc.M801406200

Achard, P., Herr, A., Baulcombe, D. C., & Harberd, N. P. (2004). Modulation of floral development by a gibberellin-regulated microRNA. *Development*, *131*(14), 3357–3365. https://doi.org/10.1242/dev.01206

Addo-Quaye, C., Eshoo, T. W., Bartel, D. P., & Axtell, M. J. (2008). Endogenous siRNA and miRNA Targets Identified by Sequencing of the Arabidopsis Degradome. *Current Biology*, *18*(10), 758–762. https://doi.org/10.1016/j.cub.2008.04.042

Addo-Quaye, C., Miller, W., & Axtell, M. J. (2009). CleaveLand: A pipeline for using degradome data to find cleaved small RNA targets. *Bioinformatics*. https://doi.org/10.1093/bioinformatics/btn604

Adenot, X., Elmayan, T., Lauressergues, D., Boutet, S., Bouché, N., Gasciolli, V., & Vaucheret, H. (2006). DRB4-Dependent TAS3 trans-Acting siRNAs Control Leaf Morphology through AGO7. *Current Biology*. https://doi.org/10.1016/j.cub.2006.03.035

Allen, E., Xie, Z., Gustafson, A. M., & Carrington, J. C. (2005). microRNA-directed phasing during trans-acting siRNA biogenesis in plants. *Cell*. https://doi.org/10.1016/j.cell.2005.04.004

Allen, R. S., Li, J., Alonso-Peral, M. M., White, R. G., Gubler, F., & Millar, A. A. (2010). MicroR159 regulation of most conserved targets in Arabidopsis has negligible phenotypic effects. *Silence*, *1*(1), 18. https://doi.org/10.1186/1758-907X-1-18

Allen, R. S., Li, J., Stahle, M. I., Dubroué, A., Gubler, F., & Millar, A. A. (2007). Genetic analysis reveals functional redundancy and the major target genes of the Arabidopsis miR159 family. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(41), 16371–16376. https://doi.org/10.1073/pnas.0707653104

Aukerman, M. J., & Sakai, H. (2003). Regulation of Flowering Time and Floral Organ Identity by a MicroRNA and Its. *The Plant Cell*.

Axtell, M. J. (2008). Evolution of microRNAs and their targets: Are all microRNAs biologically relevant? *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, *1779*(11), 725–734. https://doi.org/10.1016/j.bbagrm.2008.02.007

Axtell, M. J., & Meyers, B. C. (2018). Revisiting Criteria for Plant MicroRNA Annotation in the Era of Big Data. *The Plant Cell*, *30*(2), 272–284. https://doi.org/10.1105/tpc.17.00851

Axtell, M. J., Snyder, J. A., & Bartel, D. P. (2007). Common Functions for Diverse Small RNAs of Land Plants. *The Plant Cell*, *19*(6), 1750–1769. https://doi.org/10.1105/tpc.107.051706

Beauclair, L., Yu, A., & Bouché, N. (2010). MicroRNA-directed cleavage and translational repression of the copper chaperone for superoxide dismutase mRNA in Arabidopsis. *Plant Journal*. https://doi.org/10.1111/j.1365-313X.2010.04162.x

Bohmert, K., Camus, I., Bellini, C., Bouchez, D., Caboche, M., & Benning, C. (1998). AGO1 defines a novel locus of Arabidopsis controlling leaf development. *The EMBO Journal*, *17*(1), 170–180. https://doi.org/https://doi.org/10.1093/emboj/17.1.170

Bologna, N. G., Iselin, R., Abriata, L. A., Sarazin, A., Pumplin, N., Jay, F., Grentzinger, T., Dal Peraro, M., & Voinnet, O. (2018). Nucleo-cytosolic Shuttling of ARGONAUTE1 Prompts a Revised Model of the Plant MicroRNA Pathway. *Molecular Cell*. https://doi.org/10.1016/j.molcel.2018.01.007

Bologna, N. G., & Voinnet, O. (2014). The diversity, biogenesis, and activities of endogenous silencing small RNAs in Arabidopsis. In *Annual Review of Plant Biology*. https://doi.org/10.1146/annurev-arplant-050213-035728

Bonnet, E., He, Y., Billiau, K., & Van de Peer, Y. (2010). TAPIR, a web server for the prediction of plant microRNA targets, including target mimics. *Bioinformatics*, *26*(12), 1566–1568. https://doi.org/10.1093/bioinformatics/btq233

Borges, F., & Martienssen, R. A. (2015). The expanding world of small RNAs in plants. In *Nature Reviews Molecular Cell Biology*. https://doi.org/10.1038/nrm4085

Brodersen, P., Sakvarelidze-Achard, L., Bruun-Rasmussen, M., Dunoyer, P., Yamamoto, Y. Y., Sieburth, L., & Voinnet, O. (2008). Widespread translational inhibition by plant miRNAs and siRNAs. *Science*. https://doi.org/10.1126/science.1159151

Brousse, C., Liu, Q., Beauclair, L., Deremetz, A., Axtell, M. J., & Bouché, N. (2014). A non-canonical plant microRNA target site. *Nucleic Acids Research*, *42*(8), 5270–5279. https://doi.org/10.1093/nar/gku157

Casadevall, R., Rodriguez, R. E., Debernardi, J. M., Palatnik, J. F., & Casati, P. (2013). Repression of Growth Regulating Factors by the MicroRNA396 Inhibits Cell Proliferation by UV-B Radiation in Arabidopsis Leaves. *The Plant Cell*, *25*(9), 3570–3583. https://doi.org/10.1105/tpc.113.117473

Chávez Montes, R. A., De Fátima Rosas-Cárdenas, F., De Paoli, E., Accerbi, M., Rymarquis, L. A., Mahalingam, G., Marsch-Martínez, N., Meyers, B. C., Green, P. J., & De Folter, S. (2014). Sample sequencing of vascular plants demonstrates widespread conservation and divergence of microRNAs. *Nature Communications*. https://doi.org/10.1038/ncomms4722

Chen, K., Wang, Y., Zhang, R., Zhang, H., & Gao, C. (2019). CRISPR/Cas Genome Editing and Precision Plant Breeding in Agriculture. In *Annual Review of Plant Biology*. https://doi.org/10.1146/annurev-arplant-050718-100049

Chiou, T.-J., Aung, K., Lin, S.-I., Wu, C.-C., Chiang, S.-F., & Su, C. (2006). Regulation of Phosphate Homeostasis by MicroRNA in Arabidopsis. *The Plant Cell*, *18*(2), 412–421. https://doi.org/10.1105/tpc.105.038943

Chorostecki, U., Crosa, V. A., Lodeyro, A. F., Bologna, N. G., Martin, A. P., Carrillo, N., Schommer, C., & Palatnik, J. F. (2012). Identification of new microRNA-regulated genes by conserved targeting in plant species. *Nucleic Acids Research*, *40*(18), 8893–8904. https://doi.org/10.1093/nar/gks625

Chorostecki, U., & Palatnik, J. F. (2014). comTAR: a web tool for the prediction and characterization of conserved microRNA targets in plants. *Bioinformatics*, *30*(14), 2066–2067. https://doi.org/10.1093/bioinformatics/btu147

Clarke, J. H., Tack, D., Findlay, K., Van Montagu, M., & Van Lijsebettens, M. (1999). The SERRATE locus controls the formation of the early juvenile leaves and phase length in Arabidopsis. *Plant Journal*. https://doi.org/10.1046/j.1365-313X.1999.00623.x

Cui, J., You, C., & Chen, X. (2017). The evolution of microRNAs in plants. In *Current Opinion in Plant Biology*. https://doi.org/10.1016/j.pbi.2016.11.006

Cuperus, J. T., Fahlgren, N., & Carrington, J. C. (2011). Evolution and functional diversification of MIRNA genes. *The Plant Cell*, *23*(2), 431–442. https://doi.org/10.1105/tpc.110.082784

Dai, X., & Zhao, P. X. (2011). PsRNATarget: A plant small RNA target analysis server. *Nucleic Acids Research*, *39*(SUPPL. 2), 155–159. https://doi.org/10.1093/nar/gkr319

Dai, X., Zhuang, Z., & Zhao, P. X. (2018). PsRNATarget: A plant small RNA target analysis server (2017 release). *Nucleic Acids Research*, *46*(W1), W49–W54. https://doi.org/10.1093/nar/gky316

Dang, Y., Yang, Q., Xue, Z., & Liu, Y. (2011). RNA interference in fungi: Pathways, functions, and applications. In *Eukaryotic Cell*. https://doi.org/10.1128/EC.05109-11

Dang, Y., Yang, Q., Xue, Z., & Liu, Y. (2011). RNA interference in fungi: Pathways, functions, and applications. In *Eukaryotic Cell*. https://doi.org/10.1128/EC.05109-11

Dang, Y., Yang, Q., Xue, Z., & Liu, Y. (2011). RNA interference in fungi: Pathways, functions, and applications. In *Eukaryotic Cell*. https://doi.org/10.1128/EC.05109-11

Debernardi, J. M., Rodriguez, R. E., Mecchia, M. A., & Palatnik, J. F. (2012). Functional specialization of the plant miR396 regulatory network through distinct microRNA-target interactions. *PLoS Genetics*. https://doi.org/10.1371/journal.pgen.1002419

Deng, F., Zeng, F., Shen, Q., Abbas, A., Cheng, J., Jiang, W., Chen, G., Shah, A. N., Holford, P., Tanveer, M., Zhang, D., & Chen, Z. H. (2022). Molecular evolution and functional modification of plant miRNAs with CRISPR. In *Trends in Plant Science*. https://doi.org/10.1016/j.tplants.2022.01.009

Deveson, I., Li, J., & Millar, A. A. (2013). MicroRNAs with analogous target complementarities perform with highly variable efficacies in Arabidopsis. *FEBS Letters*, *587*(22), 3703–3708. https://doi.org/10.1016/j.febslet.2013.09.037

Dugas, D. V., & Bartel, B. (2008). Sucrose induction of Arabidopsis miR398 represses two Cu/Zn superoxide dismutases. *Plant Molecular Biology*. https://doi.org/10.1007/s11103-008-9329-1

Fahlgren, N., Howell, M. D., Kasschau, K. D., Chapman, E. J., Sullivan, C. M., Cumbie, J. S., Givan, S. A., Law, T. F., Grant, S. R., Dangl, J. L., & Carrington, J. C. (2007). High-throughput sequencing of Arabidopsis microRNAs: Evidence for frequent birth and death of MIRNA genes. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0000219

Fahlgren, N., Montgomery, T. A., Howell, M. D., Allen, E., Dvorak, S. K., Alexander, A. L., & Carrington, J. C. (2006). Regulation of AUXIN RESPONSE FACTOR3 by TAS3 ta-siRNA Affects Developmental Timing and Patterning in Arabidopsis. *Current Biology*. https://doi.org/10.1016/j.cub.2006.03.065

Fei, Y., Mao, Y., Shen, C., Wang, R., Zhang, H., & Huang, J. (2020). WPMIAS: Whole-degradome-based plant MicroRNA–Target interaction analysis server. *Bioinformatics*, *36*(6), 1937–1939. https://doi.org/10.1093/bioinformatics/btz820

Floyd, S. K., & Bowman, J. L. (2004). Ancient microRNA target sequences in plants. *Nature*. https://doi.org/10.1038/428485a

Folkes, L., Moxon, S., Woolfenden, H. C., Stocks, M. B., Szittya, G., Dalmay, T., & Moulton, V. (2012). PAREsnip: A tool for rapid genome-wide discovery of small RNA/target interactions evidenced through degradome sequencing. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gks277

Fouracre, J. P., Chen, V. J., & Scott Poethig, R. (2020). Altered meristem program1 regulates leaf identity independently of miR156-mediated translational repression. *Development (Cambridge)*. https://doi.org/10.1242/DEV.186874

Franco-Zorrilla, J. M., Valli, A., Todesco, M., Mateos, I., Puga, M. I., Rubio-Somoza, I., Leyva, A., Weigel, D., García, J. A., & Paz-Ares, J. (2007). Target mimicry provides a new mechanism for regulation of microRNA activity. *Nature Genetics*. https://doi.org/10.1038/ng2079

Gandikota, M., Birkenbihl, R. P., Höhmann, S., Cardon, G. H., Saedler, H., & Huijser, P. (2007). The miRNA156/157 recognition element in the 3′ UTR of the Arabidopsis SBP box gene SPL3 prevents early flowering by translational inhibition in seedlings. *Plant Journal*. https://doi.org/10.1111/j.1365-313X.2006.02983.x

Garcia, D. (2008). A miRacle in plant development: Role of microRNAs in cell differentiation and patterning. In *Seminars in Cell and Developmental Biology*. https://doi.org/10.1016/j.semcdb.2008.07.013

Garcia, D., Collier, S. A., Byrne, M. E., & Martienssen, R. A. (2006). Specification of Leaf Polarity in Arabidopsis via the trans-Acting siRNA Pathway. *Current Biology*. https://doi.org/10.1016/j.cub.2006.03.064

Gasciolli, V., Mallory, A. C., Bartel, D. P., & Vaucheret, H. (2005). Partially redundant functions of arabidopsis DICER-like enzymes and a role for DCL4 in producing trans-Acting siRNAs. *Current Biology*. https://doi.org/10.1016/j.cub.2005.07.024

German, M. A., Luo, S., Schroth, G., Meyers, B. C., & Green, P. J. (2009). Construction of Parallel Analysis of RNA Ends (PARE) libraries for the study of cleaved miRNA targets and the RNA degradome. *Nature Protocols*, *4*(3), 356–362. https://doi.org/10.1038/nprot.2009.8

German, M. A., Pillay, M., Jeong, D.-H., Hetawal, A., Luo, S., Janardhanan, P., Kannan, V., Rymarquis, L. A., Nobuta, K., German, R., De Paoli, E., Lu, C., Schroth, G., Meyers, B. C., & Green, P. J. (2008). Global identification of microRNA–target RNA pairs by parallel analysis of RNA ends. *Nature Biotechnology*, *26*(8), 941–946. https://doi.org/10.1038/nbt1417

Gu, W., Wang, X., Zhai, C., Xie, X., & Zhou, T. (2012). Selection on synonymous sites for increased accessibility around mirna binding sites in plants. *Molecular Biology and Evolution*. https://doi.org/10.1093/molbev/mss109

Han, M. H., Goud, S., Song, L., & Fedoroff, N. (2004). The Arabidopsis double-stranded RNA-binding protein HYL1 plays a role in microRNA-mediated gene regulation. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.0307969100

Hong, Y., Meng, J., He, X., Zhang, Y., Liu, Y., Zhang, C., Qi, H., & Luan, Y. (2021). Editing mir482b and mir482c simultaneously by crispr/cas9 enhanced tomato resistance to phytophthora infestans. *Phytopathology*. https://doi.org/10.1094/PHYTO-08-20-0360-R

Hu, J. Y., Zhou, Y., He, F., Dong, X., Liu, L. Y., Coupland, G., Turck, F., & de Meaux, J. (2014). miR824-regulated AGAMOUS-LIKE16 contributes to flowering time repression in Arabidopsis. *Plant Cell*. https://doi.org/10.1105/tpc.114.124685

Huang, J., Yang, M., & Zhang, X. (2016). The function of small RNAs in plant biotic stress response. *Journal of Integrative Plant Biology*. https://doi.org/10.1111/jipb.12463

Hunter, C., Willmann, M. R., Wu, G., Yoshikawa, M., Gutiérrez-Nava, M. de la L., & Poethig, R. S. (2006). Trans-acting siRNA-mediated repression of ETTIN and ARF4 regulates heteroblasty in Arabidopsis. *Development*. https://doi.org/10.1242/dev.02491

Ivashuta, S., Banks, I. R., Wiggins, B. E., Zhang, Y., Ziegler, T. E., Roberts, J. K., & Heck, G. R. (2011). Regulation of gene expression in plants through miRNA inactivation. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0021330

Iwakawa, H., & Tomari, Y. (2013). Molecular Insights into microRNA-Mediated Translational Repression in Plants. *Molecular Cell*, *52*(4), 591–601. https://doi.org/https://doi.org/10.1016/j.molcel.2013.10.033

Jones-Rhoades, M. W. (2012). Conservation and divergence in plant microRNAs. *Plant Molecular Biology*, *80*(1), 3–16. https://doi.org/10.1007/s11103-011-9829-2

Jones-Rhoades, M. W., & Bartel, D. P. (2004). Computational identification of plant MicroRNAs and their targets, including a stress-induced miRNA. *Molecular Cell*. https://doi.org/10.1016/j.molcel.2004.05.027

Kedde, M., van Kouwenhove, M., Zwart, W., Oude Vrielink, J. A. F., Elkon, R., & Agami, R. (2010). A Pumilio-induced RNA structure switch in p27-3’ UTR controls miR-221 and miR-222 accessibility. *Nature Cell Biology*, *12*(10), 1014–1020. https://doi.org/10.1038/ncb2105

Kidner, C. A., & Martienssen, R. A. (2004). Spatially restricted microRNA directs leaf polarity through ARGONAUTE1. *Nature*, *428*(6978), 81–84. https://doi.org/10.1038/nature02366

Kinoshita, N., Wang, H., Kasahara, H., Liu, J., MacPherson, C., Machida, Y., Kamiya, Y., Hannah, M. A., & Chua, N.-H. (2012). IAA-Ala Resistant3, an Evolutionarily Conserved Target of miR167, Mediates Arabidopsis Root Architecture Changes during High Osmotic Stress. *The Plant Cell*, *24*(9), 3590–3602. https://doi.org/10.1105/tpc.112.097006

Kozomara, A., Birgaoanu, M., & Griffiths-Jones, S. (2019). miRBase: from microRNA sequences to function. *Nucleic Acids Research*, *47*(D1), D155–D162. https://doi.org/10.1093/nar/gky1141

Kurihara, Y., Takashi, Y., & Watanabe, Y. (2006). The interaction between DCL1 and HYL1 is important for efficient and precise processing of pri-miRNA in plant microRNA biogenesis. *RNA*. https://doi.org/10.1261/rna.2146906

Kutter, C., Schöb, H., Stadler, M., Meins Jr., F., & Si-Ammour, A. (2007). MicroRNA-Mediated Regulation of Stomatal Development in Arabidopsis. *The Plant Cell*, *19*(8), 2417–2429. https://doi.org/10.1105/tpc.107.050377

Laufs, P., Peaucelle, A., Morin, H., & Traas, J. (2004). MicroRNA regulation of the CUC genes is required for boundary size control in Arabidopsis meristems. *Development*, *131*(17), 4311–4322. https://doi.org/10.1242/dev.01320

Lee, B. H., Jeon, J. O., Lee, M. M., & Kim, J. H. (2015). Genetic interaction between GROWTH-REGULATING FACTOR and CUP-SHAPED COTYLEDON in organ separation. *Plant Signaling & Behavior*, *10*(2), e988071. https://doi.org/10.4161/15592324.2014.988071

Li, F., Zheng, Q., Vandivier, L. E., Willmann, M. R., Chen, Y., & Gregory, B. D. (2012). Regulatory impact of RNA secondary structure across the arabidopsis transcriptomeW OA. *Plant Cell*. https://doi.org/10.1105/tpc.112.104232

Li, S., Liu, L., Zhuang, X., Yu, Y., Liu, X., Cui, X., Ji, L., Pan, Z., Cao, X., Mo, B., Zhang, F., Raikhel, N., Jiang, L., & Chen, X. (2013a). MicroRNAs inhibit the translation of target mRNAs on the endoplasmic reticulum in arabidopsis. *Cell*. https://doi.org/10.1016/j.cell.2013.04.005

Li, J. F., Chung, H. S., Niu, Y., Bush, J., McCormack, M., & Sheen, J. (2013b). Comprehensive protein-based artificial microRNA screens for effective gene silencing in plants. *Plant Cell*. https://doi.org/10.1105/tpc.113.112235

Li, J., Yang, Z., Yu, B., Liu, J., & Chen, X. (2005). Methylation protects miRNAs and siRNAs from a 3′-end uridylation activity in Arabidopsis. *Current Biology*. https://doi.org/10.1016/j.cub.2005.07.029

Li, J., Reichel, M., & Millar, A. A. (2014a). Determinants beyond Both Complementarity and Cleavage Govern MicroR159 Efficacy in Arabidopsis. *PLoS Genetics*. https://doi.org/10.1371/journal.pgen.1004232

Li, J., Reichel, M., Li, Y., & Millar, A. A. (2014b). The functional scope of plant microRNA-mediated silencing. *Trends in Plant Science*, *19*(12), 750–756. https://doi.org/10.1016/j.tplants.2014.08.006

Li, S., Le, B., Ma, X., Li, S., You, C., Yu, Y., Zhang, B., Liu, L., Gao, L., Shi, T., Zhao, Y., Mo, B., Cao, X., & Chen, X. (2016). Biogenesis of phased siRNAs on membrane-bound polysomes in Arabidopsis. *ELife*. https://doi.org/10.7554/elife.22750

Li, S., Xu, R., Li, A., Liu, K., Gu, L., Li, M., Zhang, H., Zhang, Y., Zhuang, S., Wang, Q., Gao, G., Li, N., Zhang, C., Li, Y., & Yu, B. (2018). SMA1, a homolog of the splicing factor Prp28, has a multifaceted role in miRNA biogenesis in Arabidopsis. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gky591

Li, W.-X., Oono, Y., Zhu, J., He, X.-J., Wu, J.-M., Iida, K., Lu, X.-Y., Cui, X., Jin, H., & Zhu, J.-K. (2008). The Arabidopsis NFYA5 Transcription Factor Is Regulated Transcriptionally and Posttranscriptionally to Promote Drought Resistance. *The Plant Cell*, *20*(8), 2238–2251. https://doi.org/10.1105/tpc.108.059444

Lian, H., Wang, L., Ma, N., Zhou, C. M., Han, L., Zhang, T. Q., & Wang, J. W. (2021). Redundant and specific roles of individual MIR172 genes in plant development. *PLoS Biology*. https://doi.org/10.1371/journal.pbio.3001044

Liang, G., He, H., Li, Y., Wang, F., & Yu, D. (2014). Molecular mechanism of microRNA396 mediating pistil development in Arabidopsis. *Plant Physiology*. https://doi.org/10.1104/pp.113.225144

Liang, G., Yang, F., & Yu, D. (2010). MicroRNA395 mediates regulation of sulfate accumulation and allocation in Arabidopsis thaliana. *The Plant Journal*, *62*(6), 1046–1057. https://doi.org/https://doi.org/10.1111/j.1365-313X.2010.04216.x

Lin, W.-Y., Lin, Y.-Y., Chiang, S.-F., Syu, C., Hsieh, L.-C., & Chiou, T.-J. (2018). Evolution of microRNA827 targeting in the plant kingdom. *New Phytologist*, *217*(4), 1712–1725. https://doi.org/https://doi.org/10.1111/nph.14938

Lindow, M., Jacobsen, A., Nygaard, S., Mang, Y., & Krogh, A. (2007). Intragenomic Matching Reveals a Huge Potential for miRNA-Mediated Regulation in Plants. *PLOS Computational Biology*, *3*(11), e238. https://doi.org/10.1371/journal.pcbi.0030238

Lindow, M., & Krogh, A. (2005). Computational evidence for hundreds of non-conserved plant microRNAs. *BMC Genomics*, *6*(1), 119. https://doi.org/10.1186/1471-2164-6-119

Liu, Q., Wang, F., & Axtell, M. J. (2014). Analysis of Complementarity Requirements for Plant MicroRNA Targeting Using a Nicotiana benthamiana Quantitative Transient Assay. *The Plant Cell*, *26*(2), 741–753. https://doi.org/10.1105/tpc.113.120972

Liu, Y., Wang, K., Li, D., Yan, J., & Zhang, W. (2017a). Enhanced Cold Tolerance and Tillering in Switchgrass (Panicum virgatum L.) by Heterologous Expression of Osa-miR393a. *Plant and Cell Physiology*, *58*(12), 2226–2240. https://doi.org/10.1093/pcp/pcx157

Liu, W. W., Meng, J., Cui, J., & Luan, Y. S. (2017b). Characterization and function of MicroRNA∗s in plants. In *Frontiers in Plant Science*. https://doi.org/10.3389/fpls.2017.02200

Llave, C., Xie, Z., Kasschau, K. D., & Carrington, J. C. (2002). Cleavage of Scarecrow-like mRNA targets directed by a class of Arabidopsis miRNA. *Science*. https://doi.org/10.1126/science.1076311

Lu, C., & Fedoroff, N. (2000). A mutation in the Arabidopsis HYL1 gene encoding a dsRNA binding protein affects responses to abscisic acid, auxin, and cytokinin. *Plant Cell*. https://doi.org/10.1105/tpc.12.12.2351

Ma, C., Burd, S., & Lers, A. (2015). MiR408 is involved in abiotic stress responses in Arabidopsis. *Plant Journal*. https://doi.org/10.1111/tpj.12999

Ma, X., Liu, C., Gu, L., Mo, B., Cao, X., & Chen, X. (2018). TarHunter, a tool for predicting conserved microRNA targets and target mimics in plants. *Bioinformatics*. https://doi.org/10.1093/bioinformatics/btx797

Machida, S., Chen, H. Y., & Adam Yuan, Y. (2011). Molecular insights into miRNA processing by Arabidopsis thaliana SERRATE. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkr428

Mallory, A. C., Bartel, D. P., & Bartel, B. (2005). MicroRNA-Directed Regulation of Arabidopsis AUXIN RESPONSE FACTOR17 Is Essential for Proper Development and Modulates Expression of Early Auxin Response Genes. *Development*, *17*(May), 1–16. https://doi.org/10.1105/tpc.105.031716.1

Mallory, A. C., Dugas, D. V, Bartel, D. P., & Bartel, B. (2004a). MicroRNA Regulation of NAC-Domain Targets Is Required for Proper Formation and Separation of Adjacent Embryonic, Vegetative, and Floral Organs. *Current Biology*, *14*(12), 1035–1046. https://doi.org/10.1016/j.cub.2004.06.022.

Mallory, A. C., Reinhart, B. J., Jones-Rhoades, M. W., Tang, G., Zamore, P. D., Barton, M. K., & Bartel, D. P. (2004b). MicroRNA control of PHABULOSA in leaf development: importance of pairing to the microRNA 5’ region. *The EMBO Journal*, *23*(16), 3356–3364. https://doi.org/10.1038/sj.emboj.7600340

Marin, E., Jouannet, V., Herz, A., Lokerse, A. S., Weijers, D., Vaucheret, H., Nussaume, L., Crespi, M. D., & Maizel, A. (2010). mir390, Arabidopsis TAS3 tasiRNAs, and their AUXIN RESPONSE FACTOR targets define an autoregulatory network quantitatively regulating lateral root growth. *Plant Cell*. https://doi.org/10.1105/tpc.109.072553

Matzke, M. A., Kanno, T., & Matzke, A. J. M. (2015). RNA-directed DNA methylation: The evolution of a complex epigenetic pathway in flowering plants. *Annual Review of Plant Biology*. https://doi.org/10.1146/annurev-arplant-043014-114633

Meng, Y., Shao, C., Gou, L., Jin, Y., & Chen, M. (2011). Construction of MicroRNA- and MicroRNA*-mediated regulatory networks in plants. *RNA Biology*, *8*(6), 1124–1148. https://doi.org/10.4161/rna.8.6.17743

Meng, Y., Shao, C., Wang, H., & Chen, M. (2012). Are all the miRBase-registered microRNAs true? *RNA Biology*. https://doi.org/10.4161/rna.19230

Miao, C., Wang, Z., Zhang, L., Yao, J., Hua, K., Liu, X., Shi, H., & Zhu, J. K. (2019). The grain yield modulator miR156 regulates seed dormancy through the gibberellin pathway in rice. *Nature Communications*. https://doi.org/10.1038/s41467-019-11830-5

Millar, A. A., & Gubler, F. (2005). The Arabidopsis GAMYB-Like Genes, MYB33 and MYB65, Are MicroRNA-Regulated Genes That Redundantly Facilitate Anther Development. *The Plant Cell*, *17*(March), 705–721. https://doi.org/10.1105/tpc.104.027920.)

Montgomery, T. A., Howell, M. D., Cuperus, J. T., Li, D., Hansen, J. E., Alexander, A. L., Chapman, E. J., Fahlgren, N., Allen, E., & Carrington, J. C. (2008). Specificity of ARGONAUTE7-miR390 Interaction and Dual Functionality in TAS3 Trans-Acting siRNA Formation. *Cell*. https://doi.org/10.1016/j.cell.2008.02.033

Morel, J.-B., Godon, C., Mourrain, P., Béclin, C., Boutet, S., Feuerbach, F., Proux, F., & Vaucheret, H. (2002). Fertile Hypomorphic ARGONAUTE (ago1) Mutants Impaired in Post-Transcriptional Gene Silencing and Virus Resistance. *The Plant Cell*, *14*(3), 629–639. https://doi.org/10.1105/tpc.010358

Navarro, L., Dunoyer, P., Jay, F., Arnold, B., Dharmasiri, N., Estelle, M., Voinnet, O., & Jones, J. D. G. (2006). A plant miRNA contributes to antibacterial resistance by repressing auxin signaling. *Science*. https://doi.org/10.1126/science.1126088

Park, W., Li, J., Song, R., Messing, J., & Chen, X. (2002). CARPEL FACTORY, a Dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in Arabidopsis thaliana. *Current Biology*. https://doi.org/10.1016/S0960-9822(02)01017-5

Peng, T., Qiao, M., Liu, H., Teotia, S., Zhang, Z., Zhao, Y., Wang, B., Zhao, D., Shi, L., Zhang, C., Le, B., Rogers, K., Gunasekara, C., Duan, H., Gu, Y., Tian, L., Nie, J., Qi, J., Meng, F., … Tang, G. (2018). A Resource for Inactivation of MicroRNAs Using Short Tandem Target Mimic Technology in Model and Crop Plants. *Molecular Plant*. https://doi.org/10.1016/j.molp.2018.09.003

Peragine, A., Yoshikawa, M., Wu, G., Albrecht, H. L., & Poethig, R. S. (2004). SGS3 and SGS2/SDE1/RDR6 are required for juvenile development and the production of trans-acting siRNAs in Arabidopsis. *Genes and Development*. https://doi.org/10.1101/gad.1231804

Pilon, M. (2017). The copper microRNAs. *New Phytologist*, *213*(3), 1030–1035. https://doi.org/https://doi.org/10.1111/nph.14244

Rajagopalan, R., Vaucheret, H., Trejo, J., & Bartel, D. P. (2006). A diverse and evolutionarily fluid set of microRNAs in Arabidopsis thaliana. *Genes and Development*. https://doi.org/10.1101/gad.1476406

Reichel, M., Li, Y., Li, J., & Millar, A. A. (2015). Inhibiting plant microRNA activity: Molecular SPONGEs, target MIMICs and STTMs all display variable efficacies against target microRNAs. *Plant Biotechnology Journal*, 1–12. https://doi.org/10.1111/pbi.12327

Reichel, M., & Millar, A. A. (2015). Specificity of plant microRNA target MIMICs: Cross-targeting of miR159 and miR319. *Journal of Plant Physiology*, *180*, 45–48. https://doi.org/https://doi.org/10.1016/j.jplph.2015.03.010

Reinhart, B. J., Weinstein, E. G., Rhoades, M. W., Bartel, B., & Bartel, D. P. (2002). MicroRNAs in plants. *Genes and Development*. https://doi.org/10.1101/gad.1004402

Reis, R. S., Hart-Smith, G., Eamens, A. L., Wilkins, M. R., & Waterhouse, P. M. (2015). Gene regulation by translational inhibition is determined by Dicer partnering proteins. *Nature Plants*. https://doi.org/10.1038/nplants.2014.27

Rhoades, M. W., Reinhart, B. J., Lim, L. P., Burge, C. B., Bartel, B., & Bartel, D. P. (2002). Prediction of Plant MicroRNA Targets. *Cell*, *110*(4), 513–520. https://doi.org/https://doi.org/10.1016/S0092-8674(02)00863-2

Robert-Seilaniantz, A., MacLean, D., Jikumaru, Y., Hill, L., Yamaguchi, S., Kamiya, Y., & Jones, J. D. G. (2011). The microRNA miR393 re-directs secondary metabolite biosynthesis away from camalexin and towards glucosinolates. *Plant Journal*. https://doi.org/10.1111/j.1365-313X.2011.04591.x

Schwab, R., Palatnik, J. F., Riester, M., Schommer, C., Schmid, M., & Weigel, D. (2005). Specific effects of microRNAs on the plant transcriptome. *Developmental Cell*, *8*(4), 517–527. https://doi.org/10.1016/j.devcel.2005.01.018

Šečić, E., Kogel, K. H., & Ladera-Carmona, M. J. (2021). Biotic stress-associated microRNA families in plants. In *Journal of Plant Physiology*. https://doi.org/10.1016/j.jplph.2021.153451

Shahbaz, M., & Pilon, M. (2019). Conserved Cu-MicroRNAs in Arabidopsis thaliana Function in Copper Economy under Deficiency. In *Plants* (Vol. 8, Issue 6). https://doi.org/10.3390/plants8060141

Smith, L. M., Burbano, H. A., Wang, X., Fitz, J., Wang, G., Ural-Blimke, Y., & Weigel, D. (2015). Rapid divergence and high diversity of miRNAs and miRNA targets in the Camelineae. *Plant Journal*. https://doi.org/10.1111/tpj.12754

Song, J. B., Gao, S., Wang, Y., Li, B. W., Zhang, Y. L., & Yang, Z. M. (2016). miR394 and its target gene LCR are involved in cold stress response in Arabidopsis. *Plant Gene*, *5*, 56–64. https://doi.org/https://doi.org/10.1016/j.plgene.2015.12.001

Song, X., Li, Y., Cao, X., & Qi, Y. (2019). MicroRNAs and Their Regulatory Roles in Plant-Environment Interactions. In *Annual Review of Plant Biology*. https://doi.org/10.1146/annurev-arplant-050718-100334

Souret, F. F., Kastenmayer, J. P., & Green, P. J. (2004). AtXRN4 degrades mRNA in Arabidopsis and its substrates include selected miRNA targets. *Molecular Cell*. https://doi.org/10.1016/j.molcel.2004.06.006

Sun, Y.-H., Lu, S., Shi, R., & Chiang, V. L. (2011). *Computational Prediction of Plant miRNA Targets.* (H. Kodama & A. Komamine (Eds.); pp. 175–186). Humana Press. https://doi.org/10.1007/978-1

Sun, Y.-H., Lu, S., Shi, R., & Chiang, V. L. (2011). *Computational Prediction of Plant miRNA Targets - RNAi and Plant Gene Function Analysis: Methods and Protocols* (H. Kodama & A. Komamine (Eds.); pp. 175–186). Humana Press. https://doi.org/10.1007/978-1-61779-123-9_12

Sunkar, R., Kapoor, A., & Zhu, J.-K. (2006). Posttranscriptional Induction of Two Cu/Zn Superoxide Dismutase Genes in Arabidopsis Is Mediated by Downregulation of miR398 and Important for Oxidative Stress Tolerance. *The Plant Cell*, *18*(8), 2051–2065. https://doi.org/10.1105/tpc.106.041673

Sunkar, R., Li, Y. F., & Jagadeeswaran, G. (2012). Functions of microRNAs in plant stress responses. In *Trends in Plant Science*. https://doi.org/10.1016/j.tplants.2012.01.010

Szaker, H. M., Darkó, É., Medzihradszky, A., Janda, T., Liu, H., Charng, Y., & Csorba, T. (2019). miR824/AGAMOUS-LIKE16 Module Integrates Recurring Environmental Heat Stress Changes to Fine-Tune Poststress Development. In *Frontiers in Plant Science* (Vol. 10, p. 1454). https://www.frontiersin.org/article/10.3389/fpls.2019.01454

Tanaka, N., Itoh, H., Sentoku, N., Kojima, M., Sakakibara, H., Izawa, T., Itoh, J. I., & Nagato, Y. (2011). The COP1 ortholog PPS regulates the juvenile-adult and vegetative-reproductive phase changes in rice. *Plant Cell*. https://doi.org/10.1105/tpc.111.083436

Taylor, R. S., Tarver, J. E., Foroozani, A., & Donoghue, P. C. J. (2017). *Insights & Perspectives MicroRNA annotation of plant genomes À Do it right or not at all*. *1600113*, 1–6. https://doi.org/10.1002/bies.201600113

Taylor, R. S., Tarver, J. E., Hiscock, S. J., & Donoghue, P. C. J. (2014). Evolutionary history of plant microRNAs. *Trends in Plant Science*, *19*(3), 175–182. https://doi.org/10.1016/j.tplants.2013.11.008

Thiebaut, F., Rojas, C. A., Almeida, K. L., Grativol, C., Domiciano, G. C., Lamb, C. R., Engler, J., Hemerly, A. S., & Ferreira, P. C. (2012). Regulation of miR319 during cold stress in sugarcane. *Plant, Cell & Environment*, *35*(3), 502–512. https://doi.org/https://doi.org/10.1111/j.1365-3040.2011.02430.x

Todesco, M., Rubio-Somoza, I., Paz-Ares, J., & Weigel, D. (2010). A collection of target mimics for comprehensive analysis of MicroRNA function in Arabidopsis thaliana. *PLoS Genetics*, *6*(7), 1–10. https://doi.org/10.1371/journal.pgen.1001031

Vaucheret, H., Vazquez, F., Crété, P., & Bartel, D. P. (2004). The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. *Genes & Development*, *18*(10), 1187–1197. https://doi.org/10.1101/gad.1201404

Vazquez, F., Gasciolli, V., Crété, P., & Vaucheret, H. (2004). The Nuclear dsRNA Binding Protein HYL1 Is Required for MicroRNA Accumulation and Plant Development, but Not Posttranscriptional Transgene Silencing. *Current Biology*, *14*(4), 346–351. https://doi.org/https://doi.org/10.1016/j.cub.2004.01.035

Voinnet, O. (2009). Origin, Biogenesis, and Activity of Plant MicroRNAs. In *Cell*. https://doi.org/10.1016/j.cell.2009.01.046

Waititu, J. K., Zhang, C., Liu, J., & Wang, H. (2020). Plant non-coding rnas: Origin, biogenesis, mode of action and their roles in abiotic stress. In *International Journal of Molecular Sciences*. https://doi.org/10.3390/ijms21218401

Wang, J. W., Czech, B., & Weigel, D. (2009). miR156-Regulated SPL Transcription Factors Define an Endogenous Flowering Pathway in Arabidopsis thaliana. *Cell*. https://doi.org/10.1016/j.cell.2009.06.014

Wang, J., Mei, J., & Ren, G. (2019). Plant microRNAs: Biogenesis, homeostasis, and degradation. In *Frontiers in Plant Science*. https://doi.org/10.3389/fpls.2019.00360

Wang, Z., Ma, Z., Castillo-González, C., Sun, D., Li, Y., Yu, B., Zhao, B., Li, P., & Zhang, X. (2018). SWI2/SNF2 ATPase CHR2 remodels pri-miRNAs via Serrate to impede miRNA production. *Nature*. https://doi.org/10.1038/s41586-018-0135-x

Williams, L., Carles, C. C., Osmont, K. S., & Fletcher, J. C. (2005). A database analysis method identifies an endogenous trans-acting short-interfering RNA that targets the Arabidopsis ARF2, ARF3, and ARF4 genes. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.0504029102

Wong, G., Alonso-Peral, M., Li, B., Li, J., & Millar, A. A. (2018). MicroRNA MIMIC binding sites: Minor flanking nucleotide alterations can strongly impact MIMIC silencing efficacy in Arabidopsis. *Plant Direct*. https://doi.org/10.1002/pld3.88

Wu, G., & Poethig, R. S. (2006). Temporal regulation of shoot development in Arabidopsis thaliana by miRr156 and its target SPL3. *Development*, *133*(18), 3539–3547. https://doi.org/10.1242/dev.02521

Wu, H., Li, B., Iwakawa, H. oki, Pan, Y., Tang, X., Ling-hu, Q., Liu, Y., Sheng, S., Feng, L., Zhang, H., Zhang, X., Tang, Z., Xia, X., Zhai, J., & Guo, H. (2020). Plant 22-nt siRNAs mediate translational repression and stress adaptation. *Nature*. https://doi.org/10.1038/s41586-020-2231-y

Wu, M. F., Tian, Q., & Reed, J. W. (2006). Arabidopis microRNA167 controls patterns of ARF6 and ARF8 expression, and regulates both female and male reproduction. *Development*. https://doi.org/10.1242/dev.02602

Xia, R., Xu, J., & Meyers, B. C. (2017). The emergence, evolution, and diversification of the miR390-TAS3-ARF pathway in land plants. *Plant Cell*. https://doi.org/10.1105/tpc.17.00185

Xie, Z., Kasschau, K. D., & Carrington, J. C. (2003). Negative Feedback Regulation of Dicer-Like1 in Arabidopsis by microRNA-Guided mRNA Degradation. *Current Biology*, *13*(9), 784–789. https://doi.org/https://doi.org/10.1016/S0960-9822(03)00281-1

Xu, H., Li, Y., Zhang, K., Li, M., Fu, S., Tian, Y., Qin, T., Li, X., Zhong, Y., & Liao, H. (2021). miR169c-NFYA-C-ENOD40 modulates nitrogen inhibitory effects in soybean nodulation. *New Phytologist*. https://doi.org/10.1111/nph.17115

Xue, X. Y., Zhao, B., Chao, L. M., Chen, D. Y., Cui, W. R., Mao, Y. B., Wang, L. J., & Chen, X. Y. (2014). Interaction between Two Timing MicroRNAs Controls Trichome Distribution in Arabidopsis. *PLoS Genetics*. https://doi.org/10.1371/journal.pgen.1004266

Yan, J., Gu, Y., Jia, X., Kang, W., Pan, S., Tang, X., Chen, X., & Tang, G. (2012). Effective small RNA destruction by the expression of a short tandem target mimic in Arabidopsis. *Plant Cell*. https://doi.org/10.1105/tpc.111.094144

Yang, C., Li, D., Mao, D., Liu, X., Ji, C., Li, X., Zhao, X., Cheng, Z., Chen, C., & Zhu, L. (2013). Overexpression of microRNA319 impacts leaf morphogenesis and leads to enhanced cold tolerance in rice (Oryza sativa L.). *Plant, Cell & Environment*, *36*(12), 2207–2218. https://doi.org/https://doi.org/10.1111/pce.12130

Yang, L., Wu, G., & Poethig, R. S. (2012). Mutations in the GW-repeat protein SUO reveal a developmental function for microRNA-mediated translational repression in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.1114673109

Yang, M., Woolfenden, H. C., Zhang, Y., Fang, X., Liu, Q., Vigh, M. L., Cheema, J., Yang, X., Norris, M., Yu, S., Carbonell, A., Brodersen, P., Wang, J., & Ding, Y. (2020). Intact RNA structurome reveals mRNA structure-mediated regulation of miRNA cleavage in vivo. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkaa577

Yang, X., Ren, W., Zhao, Q., Zhang, P., Wu, F., & He, Y. (2014). Homodimerization of HYL1 ensures the correct selection of cleavage sites in primary miRNA. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gku907

Yang, X., You, C., Wang, X., Gao, L., Mo, B., Liu, L., & Chen, X. (2021). Widespread occurrence of microRNA-mediated target cleavage on membrane-bound polysomes. *Genome Biology*. https://doi.org/10.1186/s13059-020-02242-6

You, C., Cui, J., Wang, H., Qi, X., Kuo, L. Y., Ma, H., Gao, L., Mo, B., & Chen, X. (2017). Conservation and divergence of small RNA pathways and microRNAs in land plants. *Genome Biology*. https://doi.org/10.1186/s13059-017-1291-2

Yu, B., Yang, Z., Li, J., Minakhina, S., Yang, M., Padgett, R. W., Steward, R., & Chen, X. (2005). Methylation as a crucial step in plant microRNA biogenesis. *Science*. https://doi.org/10.1126/science.1107130

Yu, Y., Jia, T., & Chen, X. (2017). The 'how' and 'where' of plant microRNAs. In *New Phytologist*. https://doi.org/10.1111/nph.14834

Zhang, T. Q., Wang, J. W., & Zhou, C. M. (2015). The role of miR156 in developmental transitions in Nicotiana tabacum. *Science China Life Sciences*. https://doi.org/10.1007/s11427-015-4808-5

Zhang, Z., Guo, X., Ge, C., Ma, Z., Jiang, M., Li, T., Koiwa, H., Yang, S. W., & Zhang, X. (2017a). KETCH1 imports HYL1 to nucleus for miRNA biogenesis in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.1619755114

Zhang, Z., Hu, F., Sung, M. W., Shu, C., Castillo-González, C., Koiwa, H., Tang, G., Dickman, M., Li, P., & Zhang, X. (2017b). RISC-interacting clearing 3'- 5' exoribonucleases (RICES) degrade uridylated cleavage fragments to maintain functional RISC in Arabidopsis thaliana. *ELife*. https://doi.org/10.7554/eLife.24466

Zhang, H., Zhang, J., Yan, J., Gou, F., Mao, Y., Tang, G., Botella, J. R., & Zhu, J. K. (2017c). Short tandem target mimic rice lines uncover functions of miRNAs in regulating important agronomic traits. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.1703752114

Zhao, M., Ding, H., Zhu, J.-K., Zhang, F., & Li, W.-X. (2011). Involvement of miR169 in the nitrogen-starvation responses in Arabidopsis. *New Phytologist*, *190*(4), 906–915. https://doi.org/https://doi.org/10.1111/j.1469-8137.2011.03647.x

Zhao, Y., Lin, S., Qiu, Z., Cao, D., Wen, J., Deng, X., Wang, X., Lin, J., & Li, X. (2015). MicroRNA857 Is Involved in the Regulation of Secondary Growth of Vascular Tissues in Arabidopsis. *Plant Physiology*, *169*(4), 2539–2552. https://doi.org/10.1104/pp.15.01011

Zheng, Y., Li, Y. F., Sunkar, R., & Zhang, W. (2012). SeqTar: An effective method for identifying microRNA guided cleavage sites from degradome of polyadenylated transcripts in plants. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkr1092

Zheng, Z., Reichel, M., Deveson, I., Wong, G., Li, J., & Millar, A. A. (2017). Target RNA Secondary Structure Is a Major Determinant of miR159 Efficacy. *Plant Physiology*, *174*(3), 1764–1778. https://doi.org/10.1104/pp.16.01898

Chapter 2

TRUEE; a bioinformatic pipeline to define the functional miRNA targetome of Arabidopsis

Abbreviations

ARF6 – AUXIN RESPONSE FACTOR 6

Cat 1-4 – Category 1-4

Cat Score – category score

CIB4 – CRY2-INTERACTING BHLH4

CSD1 – COPPER/ZINC SUPEROXIDE DISMUTASE 1 (*CSD1*)

CYP38 – CYCLOPHILIN 38

GRF – GROWTH REGULATORY FACTOR

HE – high evidence

LE – low evidence

MAFFT – Multiple Alignment using Fast Fourier Transform

Max Cat – Max Category

miRNAs – MicroRNAs

MTIs – miRNA/tasiRNA-target interactions

nt – nucleotide

PGY1 – PIGGYBACK1

PORC – PROTOCHLOROPHYLLIDE OXIDOREDUCTASE C

PPR1 – PENTATRICOPEPTIDE REPEAT1

RANBP1 – RNA BINDING PROTEIN 1

RAP2.12 – RELATED TO AP2 12

RISC – RNA Induced Silencing Complex

RPF3 – RNA PROCESSING FACTOR3

SPL – SQUAMOSA PROMOTER-BINDING PROTEIN LIKE

sRNAs – small RNAs

T-plots – Target-plots

tasiRNA – trans-acting siRNA

TCP – TEOSINTE BRANCHED1, CYCLOIDEA, and PROLIFERATING CELL NUCLEAR ANTIGEN BINDING FACTOR

TP10M – transcript per 10 million

TRUEE – Targets Ranked Using Experimental Evidence

VAT – Validated Arabidopsis Target

WPMIAS – Whole-Degradome-based Plant MicroRNA-Target Interaction Analysis Server

Abstract

Central to plant microRNA (miRNA) biology is the identification of functional miRNA-target interactions (MTIs). However, the complementarity basis of bioinformatic target prediction results in mostly false positives, and the degree of complementarity does not equate with regulation. Here, we develop a bioinformatic workflow named TRUEE (Targets Ranked Using Experimental Evidence) that ranks MTIs on the extent to which they are subjected to miRNA-mediated cleavage. It sorts predicted targets into high (HE) and low evidence (LE) groupings based on the frequency and strength of miRNA-guided cleavage degradome signals across multiple degradome experiments. From this, each target is assigned a numerical value, termed a Category Score, ranking the extent to which it is subjected to miRNA-mediated cleavage. As a proof-of-concept, the 428 Arabidopsis miRNAs annotated in miRBase were processed through the TRUEE pipeline to determine the miRNA 'targetome'. The majority of high-ranking Category Score targets corresponded to highly conserved MTIs, validating the workflow. Very few Arabidopsis-specific, Brassicaceae-specific, or Conserved-passenger miRNAs had HE targets with high Category Scores. In total, only several hundred MTIs were found to have Category Scores characteristic of currently known physiologically significance MTIs. Although non-exhaustive, clearly the number of functional MTIs is much narrower than many studies claim. Therefore, using TRUEE to numerically rank targets directly on experimental evidence has given insights into the scope of the functional miRNA targetome of Arabidopsis.

2.1 Introduction

MicroRNAs (miRNAs) are short non-coding RNAs of approximately 20-22 nt in length which guide the RNA Induced Silencing Complex (RISC) to repress target mRNAs via transcript cleavage and/or translational repression. Given that a high degree of complementarity between a plant miRNA-target pair is necessary for a strong repression (Schwab et al., 2005; Liu et al., 2014), numerous bioinformatic target prediction programs based on mismatch scoring schemas have been developed (Bonnet et al., 2010; Dai & Zhao, 2011; Sun et al., 2011). These scoring schema consider the positions of mismatches, weightings for different mismatches (G:U pairs) and potential miRNA binding-site accessibility (Mallory et al., 2004; Allen et al., 2005; Schwab et al., 2005; Bonnet et al., 2010; Dai & Zhao, 2011; Sun et al., 2011). As further studies experimentally identified miRNA-target pairs with complementarity that would not be detected by these initial scoring schemas (Zheng et al., 2012; Brousse et al., 2014), this has justified relaxing complementarity requirements of the bioinformatic prediction of miRNA targets. For example, in an updated version of the most widely cited miRNA-target prediction tool, psRNATarget (version 2), the default parameter relating to complementarity (expectation score) was relaxed from 3 to 5 (Dai et al., 2018). Although this improved the prediction (or recall) of 143 of 147 experimentally validated Arabidopsis targets, there were almost 10,000 predicted targets in the bioinformatic output (Dai et al., 2018). Therefore, the output is overwhelmed with likely false positives.

It has also become evident that miRNA-target complementarity does not correlate with a functional miRNA-mediated regulatory outcome. For example, of a family of seven Arabidopsis *GAMYB-like* genes that contained analogous conserved miR159-binding sites, only two genes were found to be strongly miR159-regulated (Allen et al., 2007; Zheng et al., 2017). This, with the myriad of potential false positives, and the inability to rank targets on complementarity, highlights the limitations of identifying the cohort of functional plant miRNA-target genes using bioinformatics alone, and the need to develop a miRNA target prediction scoring schema independent of miRNA target binding site complementarity.

Degradome sequencing has been used to experimentally compliment bioinformatics approaches (Addo-Quaye et al., 2008; German et al., 2008). As miRNA guide target cleavage precisely between the 10th and 11th nucleotide of the miRNA-binding site, sequencing and then mapping of the 5` ends of degraded transcripts can accurately identify miRNA-guided cleavage products. Mapping of these degradome reads to individual transcripts form target-plots (T-plots), in which the relative abundance of reads mapping precisely to the cleavage site of a potential miRNA target (cleavage tag) can be compared to all other reads on the transcript (Addo-Quaye et al., 2008; German et al., 2008). Based on the frequency of the cleavage tag relative to the other reads in a transcript, these T-plots can then be placed into four categories [Category (Cat) 1-4], indicating the most confident (Cat 1) to least confidence (Cat 4) of a target being subjected to miRNA-guided cleavage (Addo-Quaye et al., 2008). Most canonical miRNA targets are Cat 1 targets (the cleavage tag being the most abundant read), and this is considered a hallmark of a validated target (German et al., 2008). There has now been extensive degradome analysis done in many plant species, and these data are available to determine which predicted miRNA targets have degradome signatures. For example, the Whole-degradome-based Plant MicroRNA-target Interaction Analysis Server (WPMIAS) makes data from numerous publicly available degradome libraries across diverse species easily accessible (Fei et al., 2020). However, detection of a degradome signal will be reliant on isolating RNA from a tissue in which both the miRNA and target mRNA are present, so any one single degradome library will only reflect miRNA-target interactions (MTIs) in these tissues, or in plants grown under those specific conditions. Moreover, degradome analysis only detects miRNA-mediated cleavage, but not other mechanisms, such as translational repression. Furthermore, as this is a biochemical signature, detection of a degradome signature does not necessarily equate to a miRNA-target interaction of physiological significance, nor can there be an arbitrary cutoff implying that any one particular degradome signature defines that gene as a “real” miRNA target.

Adding to this uncertainty, is the identification of *bona fide* miRNAs themselves. Currently, miRBase is the go-to repository of experimentally identified miRNAs, with the latest release (v22) detailing 1000s of different miRNA sequences that have been reported across many diverse plant species (Kozomara et al., 2019). However, many publications have queried the quality and validity of these miRNA entries which are mostly user-submitted, and have suggested the greater majority of entries are potentially false positives (Taylor et al., 2014; Axtell and Meyers, 2018). Identifying high evidence miRNA targets for these miRNAs would help determine whether these miRNAs are genuine or potentially mis-annotated small RNAs (sRNAs).

This paper develops a bioinformatic workflow that attempts to address the limitations outlined above. Long lists of putative targets from complementary-based predictions (psRNATarget), are filtered using an online server (WPMIAS) in which multiple degradome libraries can be searched for corresponding cleavage tags. The workflow then assesses the frequency and strength of the degradome signatures for each predicted target, which can then be arbitrarily sorted into high and low evidence targets, as well as non-arbitrarily ranking score based on the frequency and strength of degradome signatures for each predicted target. Using Arabidopsis as a proof-of-concept, this workflow was applied to gain a better understanding of the functional scope of a plant miRNome, by obtaining an accurate estimate of the total number of MTI that have degradome signatures characteristic of known physiologically significant MTIs (i.e. MTI that when manipulated can alter a trait). We call the collection of targets the “miRNA targetome” which estimates the number of MTIs that have degradome characteristics of physiologically relevant MTIs.

2.2 Results

2.2.1 A bioinformatic workflow to facilitate the identification of high evidence miRNA targets

A workflow was developed that sorts predicted miRNA targets into groups of either high evidence (HE) or low evidence (LE) targets, and then ranks the HE targets on the strength and frequency of their T-plots across degradome experiments. This workflow has been designated the “Targets Ranked Using Experimental Evidence” (TRUEE), and combines miRbase to retrieve miRNA sequences (Kozomara et al., 2019), psRNATarget to predict miRNA targets (Dai et al., 2018), which are then subsequently used as input into WPMIAS (Fei et al., 2020) to retrieve all corresponding degradome data (Figure 2.1). Both psRNATarget and WPMIAS were chosen as they are highly accessible via user-friendly webservers, and for psRNATarget, it is the most used and cited miRNA target prediction tool. Parameters are then implemented, filtering the degradome data to distinguish HE from LE targets of miRNA-mediated regulation, and then a simple formula for ranking HE targets. Below is the description of the input parameters and the rationale for developing the TRUEE workflow.

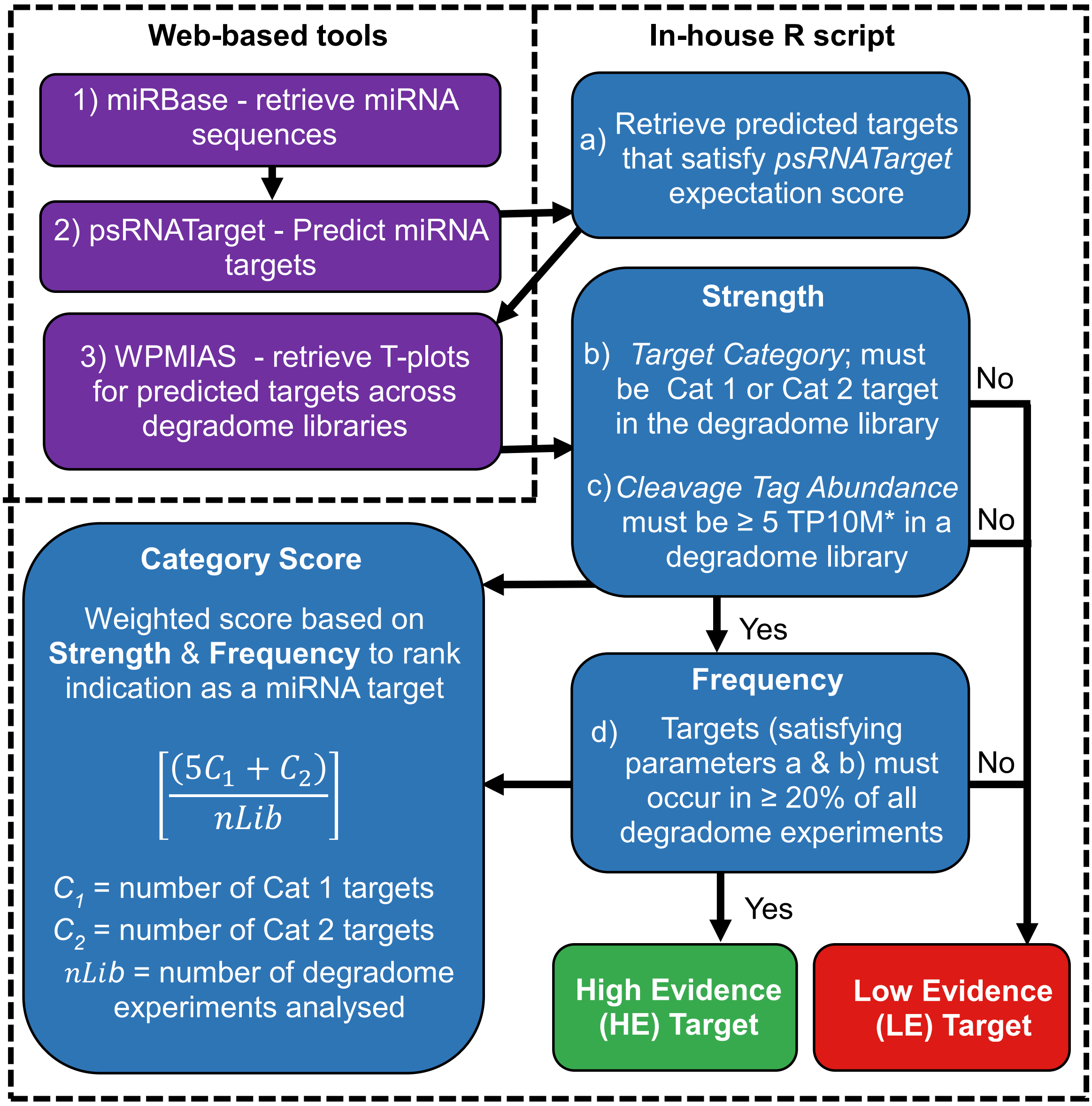


Figure 2.1. The workflow and parameters of TRUEE. Purple boxes indicate data retrieved from external web-based tools. Blue boxes indicate parameters (a) to (d) which were used to filter for HE targets and the category score (Cat Score) scoring schema. MiRNAs were retrieved from miRbase (v22) (Kozomara et al., 2019). Potential miRNA target cleavage sites were then predicted using psRNATarget (Dai et al., 2018) and predicted targets with an expectation score ≤ 3.0 or ≤ 5.0* were used for further analysis. The degradome data for these cleavage sites were then retrieved using WPMIAS (Fei et al., 2020). *TP10M means Transcript Per 10 Million.

2.2.2 An experimentally validated set of Arabidopsis miRNA targets to benchmark TRUEE parameters

To develop this workflow, the input parameters were benchmarked against a compiled set of 106 experimentally validated small RNA targets from Arabidopsis based on the literature that we have termed the “Validated Arabidopsis Target (VAT)” set. It is composed of targets of 28 miRNA families and one trans-acting siRNA (tasiRNA) family (the TAS3 phasing products, tasiARFs) and includes both widely and narrowly conserved MTIs (Table S1). To qualify as a validated target in this set, at least two independent lines of evidence from commonly used experimental approaches to identifying miRNA targets were required. This includes genetic evidence (altered mRNA/protein expression in *mirna* loss-of-function or miRNA overexpression plants, or expression of a miRNA-resistant target gene) or molecular evidence (degradome analysis, 5′-RACE cleavage assays or correlation of miRNA/target mRNA levels). The requirement of two independent lines of evidence to qualify for this list has resulted in a lower number of genes than other comparable lists in the literature (Folkes et al., 2012; Zheng et al., 2012; Srivastava et al., 2014; Dai et al., 2018; Ma et al., 2018).

2.2.3 The input parameters of TRUEE workflow

There are four parameters to consider; (a) *psRNATarget Expectation Score*; (b) *Cleavage Tag Abundance* - the number of degradome sequencing reads that coincide with the predicted cleavage site; (c) the *Target Category* - corresponding to the Cat 1-4 categories of the T-plots and (d) the *Library % Cut-off* - corresponding to the percentage of degradome libraries in which a predicted target occurs with the defined (a), (b) and (c) parameters. The optimal cut-offs for these parameters were determined via analysis of 61 Arabidopsis degradome libraries available on WPMIAS (Fei et al., 2020), from which identified targets were benchmarked against the VAT. The aim was to maximise the number of VAT targets identified, while minimizing additional targets that may represent either newly discovered targets or false positives (henceforth, potential targets).

(a) *psRNATarget Expectation Score*

The first parameter considered for TRUEE was the psRNATarget expectation score, a penalty score weighted on the number and position of mismatches between a miRNA and a predicted target gene (Dai et al., 2018). Using an expectation score too low will result in false negatives, while an expectation score too high will generate a multitude of false positives. The most recent version of psRNATarget (v2) has a default expectation score of 5, as some canonical target genes have expectation scores higher than 4 (Dai et al., 2018). As such, the expectation scores that were analysed ranged from 0 to 5. Using an expectation score of ≤ 5.0 predicted 2977 targets for the 29 miRNA/siRNA families, a greater than 28-fold increase compared to the 106 targets of the VAT set. This predicted/validated target fold difference decreased with decreasing expectation score, although fewer of the VAT set were captured (Figure 2.2A). From the analysis, an expectation score of ≤ 3 appears optimal, resulting in a relatively low-fold difference (3.5-fold), yet still included a large percentage of targets from the VAT set (89%) (Figure 2.2A). In comparison, using an expectation score any higher than 3 disproportionally increased the number of predicted targets captured (i.e. potential false positives), whereas an expectation score ≤ 2.5 failed to identify many of the VAT set (i.e. false negatives) (Figure 2.2B). For miRNAs with experimentally validated targets with an expectation scores > 3 (miR167, miR398, and miR408), the expectation score threshold was increased to ≤ 5.

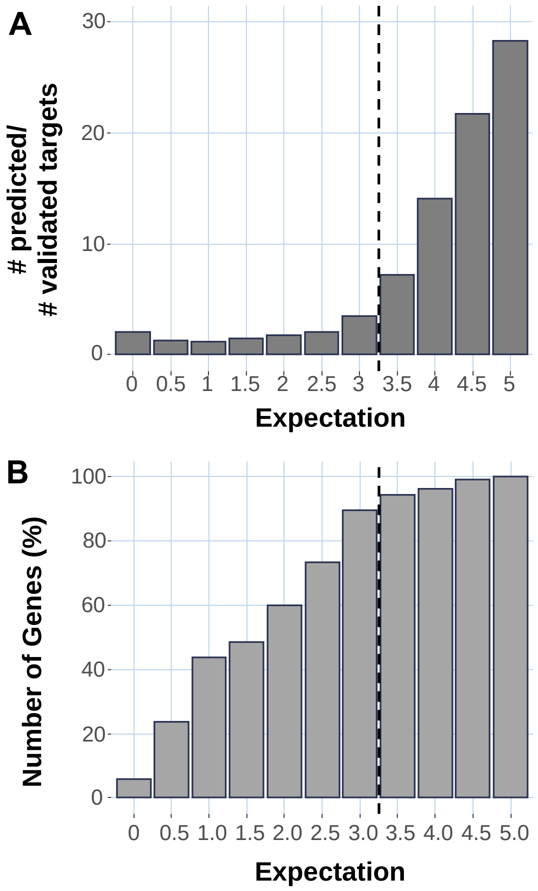


Figure 2.2. Determining the optimal psRNATarget Expectation Score cut-off. A) Fold differences of the total number of targets predicted by psRNATarget over the number of targets in the VAT set identified at each expectation score cut-off. Black numbers above each bar is the total number of predicted targets / number of validated targets for each expectation score. B) The cumulative percentage of the 106 targets of the VAT set that are retrieved at each expectation score cut-off. The red bar indicates the expectation score cut-off that was chosen for the TRUEE workflow. Total HE targets = Validated targets + potential targets

(b) *Cleavage Tag Abundance*

This parameter represents the number of cleavage tag reads for any given RNA, with the greater the read, the more confidence of miRNA-mediated regulation. Therefore, targets with a low *cleavage tag abundance* may represent fortuitous degradation events coinciding with the predicted cleavage site and thus represent a false positive. To determine an optimal value, TRUEE analysis was performed using a *Cleavage Tag Abundance* of ≥ 1, ≥ 5 and ≥ 10 when normalised to transcript per 10 million (TP10M), values that have been used in previous degradome studies (Jeong et al., 2013; Thody et al., 2020). This indicates that the degradome library for an RNA are only considered in analysis if the corresponding cleavage tag has at least 1, 5 or 10 TP10M, respectively. For this analysis, TRUEE was performed with variable *Library % cut-offs* and *Target Categories*.

A *Cleavage Tag Abundance* of ≥ 1 TP10M identified the greatest number of the VAT set (Figure 2.3A-C). At a *Library % Cut-off* of 10%, nearly all of the VAT set was identified (97%). However, the number of potential targets also almost doubled the number of the VAT set (Figure 2.3A). Furthermore, across all *Library % Cut-offs*, the number of potential targets was many fold greater compared to when the *Cleavage Tag Abundance* was set to ≥ 5 and ≥ 10 TP10M (Figure 2.3A-C).

A *cleavage tag abundance* of ≥ 5 TP10M appeared optimal. It identified a greater number of the VAT set compared to when using a setting of ≥ 10 TP10M but had a greatly reduced number of potential targets compared to when the *Cleavage Tag Abundance* was set to ≥ 1 TP10M (Figure 2.3A-C). Therefore, a *Cleavage Tag Abundance* of ≥ 5 appeared to minimize signals from potential random degradation, while maximizing identification of the VAT set.

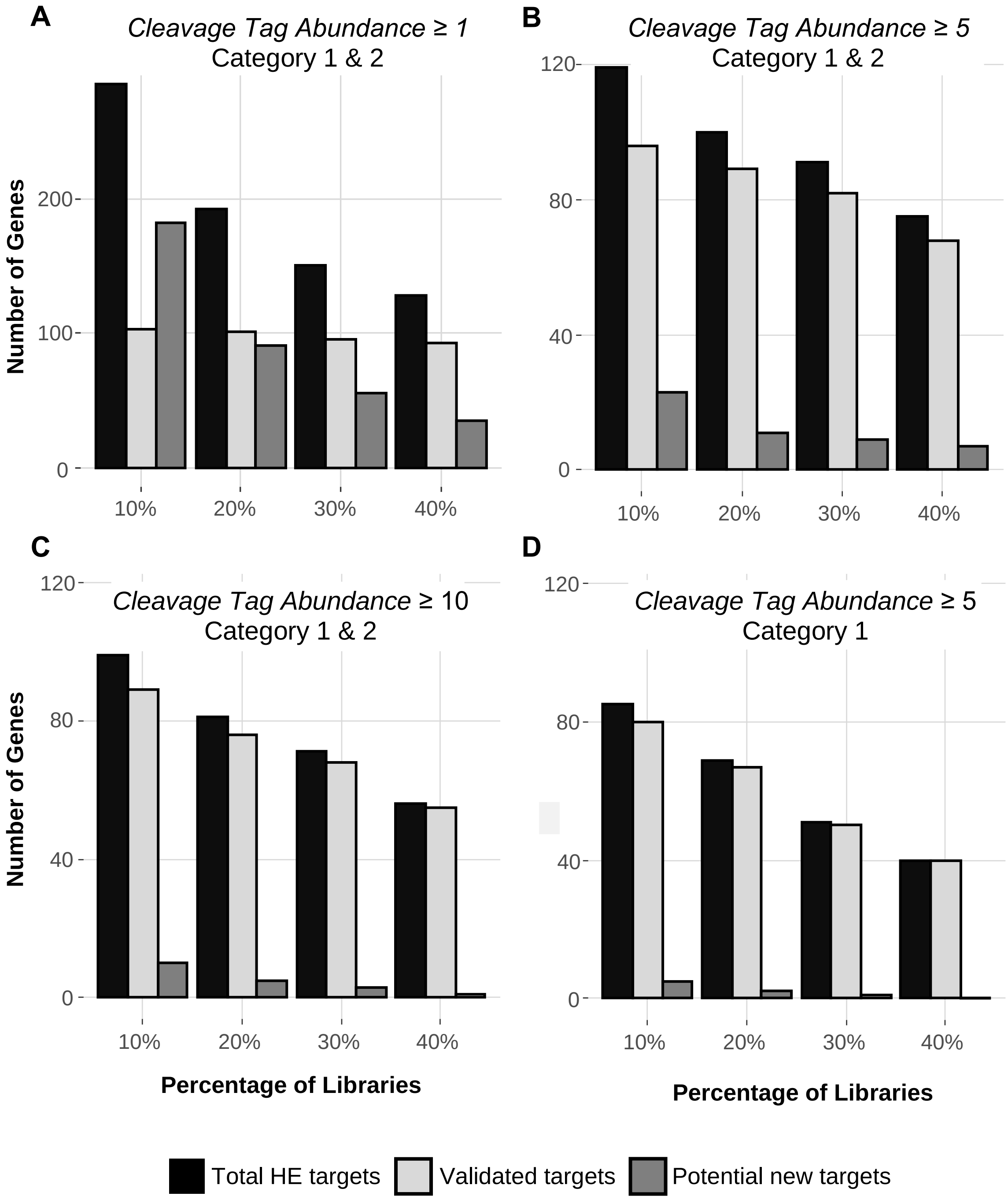


Figure 2.3. Number of genes defined as HE targets as determined by *Library % Cut-off, Cleavage Tag Abundance* or *Target Category*. 'Total targets' indicate the total number of HE targets found by TRUEE. 'Validated targets' are the HE targets found in the VAT set. 'Potential targets' are HE targets that are not found in the VAT set. Note the differences in y-axis scales.

(c) *Target Category*

For each predicted target RNA, the readout of degradome analyses are T-plots. On WPMIAS, T-plots are placed into four *Target Categories* (1-4)*,* with descending levels of confidence and so only inclusion of *Target Category* 1 and 2 targets are recommended and is set as the default (Fei et al., 2020). However, to identify the targets with greatest evidence, the stringency of TRUEE was increased by only including *Target Category* 1 targets. Results show that even at the lowest *Library % Cut-off* of 10%, only 75% of the VAT set were identified as HE targets (Figure 2.3D). This was 17 fewer targets compared to when using both *Target Category* 1 and 2 (Figure 2.3B). As only using *Target Category* 1 resulted in potentially many false negatives, for the third parameter, *Target Category* 1 and 2 were used to maximize the identification of the VAT set.

(d) *Library % Cut-off*

The parameter, *Library % Cut-off*, assesses the frequency at which a predicted target satisfies the stated parameters in all degradome libraries analysed. The greater number of libraries a predicted target occurs in, the greater evidence it has as a miRNA target. As mentioned above, the *Library % Cut-offs* were 10%, 20%, 30% and 40%. Analysis was performed on the 61 Arabidopsis degradome libraries available on WPMIAS (Fei et al., 2020).

At a *Cleavage Tag Abundance* of ≥ 5, and a *Target Category* of 1 and 2, *Library % Cut-off* was assessed (Figure 2.3B). At the *Library % Cut-off* of 10%, 97 VATs and 22 new potential targets were identified. The number of VAT set and potential targets identified decreased to 90 and 10, respectively, at a 20% *Library % Cut-off*. These values continued to decrease with increasing *Library % Cut-off*. Based on this, a *Library % Cut-off* of 20% appears optimal, as most of the VAT set was identified as HE targets, with less than 50% of additional new potential targets compared to a *Library % Cut-off* of 10%.

In conclusion, using a *Library % Cut-off* of 20%, with a *Target Category* of 1 and 2, and a *Cleavage Tag Abundance* of ≥ 5 TP10M TRUEE maximizes the identification of VAT set targets, whilst minimizing potential targets.

2.2.4 Category score (*Cat Score*); a simple scoring schema to rank HE targets

Within the HE targets identified from the above workflow, there will remain a large variation in the confidence and extent to which the retrieved targets are being subjected to miRNA-mediated degradation. Therefore, ranking these HE targets based on the strength and frequency of the target across libraries will enable a clear indication of the confidence miRNA-mediated degradation for each target. As the *Target Category* approximates the extent of which miRNA-mediated cleavage contributes to RNA degradation of a target, a scoring schema was devised which considers the number of libraries (frequency) a gene is found to be a Category 1 (C_1) or Category 2 (C_2) target with a *Cleavage Tag Abundance* of ≥ 5 (strength) (Figure 2.1). C_1 and C_2 were assigned the weighted values of 5 and 1, respectively. The heavier weighting for C_1 compared to C_2 targets was chosen considering the reduced confidence of the latter in reflecting miRNA-mediated degradation. The weighted number of libraries a gene was found to be a C_1 or C_2 target was then divided by the total number of libraries analysed (*nLib*). The category score (*Cat Score*) was calculated by:

$$Cat\ Score = \frac{(5C_1 + C_2)}{nLib}$$

This equation can give a maximum *Cat Score* = 5, which would mean the gene is a Category 1 target in all degradome libraries analysed. For such a scenario, both the miRNA and the target mRNA would need to be widely expressed so as to be detected in all degradome libraries.

Determining the *Cat Score* of the VAT set targets identified by TRUEE (Figure 2.3B) found that the *Cat Score* ranged from 4.15 to 0.12 (Table S1), enabling this ranking score to rapidly assess the extent of miRNA-mediated degradation for each HE target. Eight targets have a *Cat Score* ≥ 4, implying these MTIs are occurring strongly throughout *Arabidopsis*. Even within a family of miRNA targets, *Cat Score*s are highly variable. For instance, the *GROWTH REGUATORY FACTOR* (*GRF*) genes that are validated targets of miR396 have *Cat Score*s that vary from 4.02 (*GRF1;* At2g22840) to 0.12 (*GRF7;* At5g53660). Similarly, the *SQUAMOSA PROMOTER-BINDING PROTEIN LIKE* (*SPL*) genes that are validated targets of miR156 have *Cat Score*s that vary from 3.18 (*SPL13*; AT5G50570) to 0.33 (*SPL9*; AT2G42200), and the *TEOSINTE BRANCHED1, CYCLOIDEA, and PROLIFERATING CELL NUCLEAR ANTIGEN BINDING FACTOR* (*TCP*) genes that are validated targets of miR319 have *Cat Score*s that vary from 2.23 (*TCP4;* AT3G15030) to 0.28 (*TCP10;* AT2G31070). This enables clear identification of which paralogues with identical (or near identical) expectation scores are subjected to the strongest miRNA-mediated degradation. Additionally, having a *Cat Score* ≥ 1 would indicate that the gene must be a Category 1 target in at least one degradome library. Of the 106 VAT set, 75 (70.8%) have a *Cat Score* of ≥ 1 (Table S1). This indicates that this cut-off will identify the majority of experimentally validated miRNA targets.

2.2.5 HE targets identified by TRUEE that are not in the VAT set

In the analysis above, TRUEE identified HE targets from Arabidopsis that were not present in the VAT set, and therefore may be new targets or false positives. These targets are analysed below in terms of their *Library % Cut-off*, their *Cat Score* and their highest *Target Category* (*Maximum Category*).

To maximize the potential of identifying new targets, the *Library % Cut-off* was lowered from 20% to 10%, resulting in the identification of a total of 22 new potential targets (Table 2.1). However, the 12 additional potential targets identified at the *Library % Cut-off* of 10%, all have a very low *Cat Score* (all but two were < 0.5). This lends support to the justification of using the *Library % Cut-off* of 20% determined above. Of the 22 targets, only four had a *Cat Score* > 1, and these were in 40% of libraries. Four of these targets showed evidence that was typical of canonical miRNA targets. The highest ranked targets, *RNA PROCESSING FACTOR3* (*RPF3*) and *PENTATRICOPEPTIDE REPEAT1* (*PPR1*) are both family member homologues of genes in the VAT set with evidence of being miRNA targets, so should have likely been included in the VAT set (Howell et al., 2007; Allen et al., 2004). However, no clear previous evidence exists for the miR167 target, *RNA BINDING PROTEIN 1* (*RANBP1*) or the miR398 target, *PIGGYBACK1* (*PGY1*), both of which had a *Maximum Category* of 1 with a high *Cleavage Tag Abundance* (Figure 2.4A-B). Both T-plots were comparable to that of previously validated miR167 target, *AUXIN RESPONSE FACTOR 6* (*ARF6*), or the miR398 target, *COPPER/ZINC SUPEROXIDE DISMUTASE 1* (*CSD1*) (Figure S1E-H), suggesting an analogous degree of miRNA-mediated regulation in this library. Neither the miRNA-binding site in *RANBP1* nor *PGY1* was conserved beyond the *Brassicaceae* family (Figure S2), and so may explain why targets such as these have not been previously identified by bioinformatic tools that rely on conservation (Chorostecki et al., 2012; Ma et al., 2018).

In contrast, although *PROTOCHLOROPHYLLIDE OXIDOREDUCTASE C* (*POR C*) and *CYCLOPHILIN 38* (*CYP38*) were also found in more than 40% of libraries, they had comparatively lower *Cat Scores* (< 0.5). Additionally, their *Maximum Category* was 2, and subsequent investigation of their T-plots revealed *Cleavage Tag Abundances* to be comparable to other degradome reads mapping at many different nucleotide positions throughout the transcript (Figure 2.4C-D). This suggests the occurrence of the high *Cleavage Tag Abundance* in a high percentage of degradome libraries may be due to RNA degradation pathways other than miRNA-mediated regulation.

Despite occurring in fewer libraries than *PORC1* and *CYP38*, four additional targets, *MUSE1*, *SERINE/THREONINE-KINASE*, a *TPR* homologue and a *NAC* homologue, have greater *Cat Scores* and their *Maximum Category* was 1. Again, both *TPR* and *NAC* are family members of genes previously found to be miRNA-regulated, but for *MUSE1* and *SERINE/THREONINE-KINASE* there is no known evidence for miRNA-regulation, and both display T-plots characteristic of canonical targets (Figure 2.4E-F). This suggests that even at a *Library % Cut-off* of 10%, by considering targets with the highest *Cat Scores*, TRUEE is able to identify targets with T-plots highly indicative of miRNA-mediated cleavage. Therefore, while also considering *Library % Cut-off* and *Maximum Category*, *Cat Score* enables the ranking of targets which should be given priority for further investigation regarding potential miRNA regulation. In this regard, a *Library % Cut-off* of 10%, in addition to a *Cat Score* cut-off of ≥ 0.5, may be used as an alternate set of parameters to identify HE targets.

miRNA	**Target ID**	**Target Description**	***Library % Cut-off***				***Max Cat***	***Cat S***
			10%	**20%**	**30%**	**40%**		
miR161	AT1G62930	**RPF3, RNA Processing Factor 3**	X	X	X	X	1	2.311
miR161	AT1G06580	**PPR1, Pentatricopeptide Repeat 1**	X	X	X	X	1	1.180
miR167	AT5G58590	**RANBP1, RAN BINDING PROTEIN 1**	X	X	X	X	1	1.689
miR398	AT2G27530	**PGY1, PIGGYBACK 1**	X	X	X	X	1	1.246
miR398	AT1G03630	POR C,	X	X	X	X	2	0.492
miR408	AT3G01480	Cyclophilin 38	X	X	X	X	2	0.492
miR168	AT3G58030	**MUSE1**	X	X	X		1	0.852
miR408	AT1G68010	HPR, HYDROXYPYRUVATE REDUCTASE	X	X	X		2	0.328
miR395	AT1G50930	**Serine/Threonine-kinase**	X	X			1	0.541
miR396	AT3G19400	Cysteine proteinases superfamily protein	X	X			1	0.393
miR161	AT1G64583	**Tetratricopeptide repeat (TPR)-like**	X				1	0.721
miR164	AT3G12977	**NAC (No Apical Meristem) domain**	X				1	0.525
miR167	AT1G51760	IAR3, IAA-Alanine Resistant 3,	X				1	0.295
miR167	AT5G10550	GTE2, Global Transcription Factor E2	X				2	0.148
miR172	AT3G05530	ATS6A.2, RPT5A, TRIPLE-A ATPASE 5A	X				2	0.131
miR396	AT1G48380	HYP7, HYPOCTYL 7, ROOT HAIRLESS 1	X				1	0.262
miR396	AT1G60140	TPS10, Trehalose Phosphate Synthase	X				1	0.295
miR398	AT4G24280	cpHsc70-1, chloroplast heat shock 70-1	X				2	0.164
miR398	AT5G14550	beta-1,6-N-acetylglucosaminyltransferase	X				2	0.115
miR408	AT5G21930	PAA2, P-type ATPase of Arabidopsis 2	X				2	0.148
miR408	AT2G47900	TLP3, TUBBY LIKE PROTEIN 3	X				2	0.131
miR408	AT4G34230	CAD5, Cinnamyl Alcohol Dehydrogenase 5	X				2	0.131

Table 2.1. Analysis of identified HE targets not present in the VAT set. The *Library % Cut-off* threshold meet for each HE target is indicated by 'X'. Bolded genes indicate HE targets which were found to possess T-plots comparable to those in the VAT set. *Maximum Category* (*Max Cat*) indicates whether the highest T-plot Category found across degradome libraries is Cat 1 or Cat2 and *Cat S* is *Category Score*.

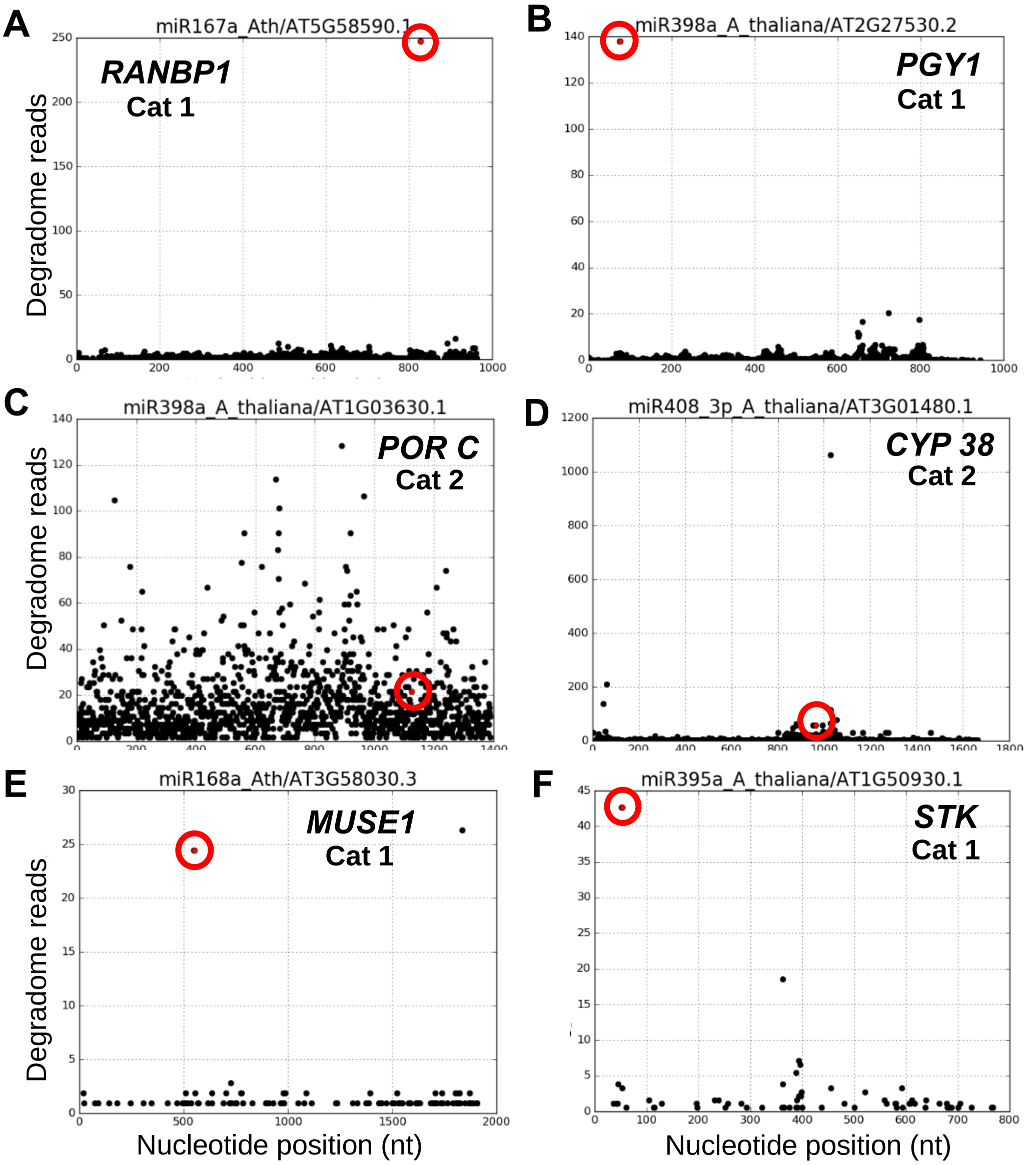


Figure 2.4. T-plots of HE targets not from the VAT set. T-plots of (A) *RNA BINDING PROTEIN 1* (*RANBP1*); (B) *PIGGYBACK1* (*PGY1*) that encodes a ribosomal protein L10aP; (C) *PROTOCHLOROPHYLLIDE OXIDOREDUCTASE C* (*POR C*); (D) *CYCLOPHILIN 38* (*CYP38*); (E) *MUSE1,* encodes a RING domain E3 ligase (F) *VASCULAR-RELATED UNKNOWN PROTEIN 2* that encodes a serine/threonine-kinase (*STK*). The T-plot from the degradome library with the highest *Maximum Category* and highest *Cleavage Tag Abundance* was used for each miRNA target. The cleavage tag is circled in red. T-plot figures were adapted from WPMIAS (Fei et al., 2020).

2.2.6 Modification of TRUEE to consider narrow spatial and temporal expression

At a *Library % Cut-off* of 20%, only 16/106 of the VAT set were not identified by TRUEE (Table S1). Several of these are known canonical miRNA targets, most of which are only regulated under specific environmental/stress conditions and so are likely being overlooked by TRUEE due to insufficient degradome libraries under the specific environmental conditions that these MTIs occur. To overcome this, analysis of select degradome libraries from a particular treatment or tissues may better detect these narrow spatial or temporal MTIs. For instance, narrowing TRUEE to only analyse root libraries finds large increases to the *Cat Score* of *SERINE/THREONINE-KINASE* (0.5 to 4.3), and a *NAC* homologue (At3g12977) (0.5 to 3.33), implying these MTIs occur preferentially in roots (Table 2.2). Therefore, by filtering which degradome libraries are analysed, TRUEE can allow the identification of more subtle MTIs, such as spatially specific MTIs.

miRNA	**Target ID**	**Target Description**	***Library % Cut-off***				***Max***	***Cat S***
			10%	**20%**	**30%**	**40%**	***Cat***	
miR161	AT1G06580	PPR1, Pentatricopeptide Repeat 1	X	X	X	X	1	4.167
miR161	AT1G62930	Tetratricopeptide repeat (TPR)-like superfamily protein	X	X	X	X	2	0.667
miR164	**AT3G12977**	**NAC (No Apical Meristem) domain**	**X**	**X**	**X**	**X**	**1**	**3.333**
miR172	AT3G05530	REGULATORY PARTICLE TRIPLE-A ATPASE 5A	X	X	X	X	2	0.5
miR395	**AT1G50930**	**Serine/Threonine-kinase**	**X**	**X**	**X**	**X**	**1**	**4.333**
miR396	AT3G19400	Cysteine proteinases superfamily protein	X	X	X	X	2	0.5
miR398	AT2G27530	PGY1, PIGGYBACK 1	X	X	X	X	2	1
miR396	AT1G60140	TPS10, TREHALOSE PHOSPHATE SYNTHASE	X	X	X		2	0.333
miR398	AT4G26230	Ribosomal protein L31e family protein	X	X	X		2	0.333
miR408	AT4G34230	CINNAMYL ALCOHOL DEHYDROGENASE 5	X	X	X		2	0.333
miR857	AT5G36880	ACS, acetyl-CoA synthetase	X	X	X		2	0.333
miR159	AT2G21600	endoplasmatic reticulum retrieval protein 1B	X				2	0.167
miR159	AT3G08850	RAPTOR1B	X				2	0.167
miR161	AT1G64583	Tetratricopeptide repeat (TPR)-like protein	X				1	0.833
miR163	AT5G38100	SABATH family methyltransferase.	X				1	0.833
miR166	AT1G07810	RNA-binding (RRM/RBD/RNP motifs) protein	X				2	0.167
miR167	AT3G07810	RNA-binding (RRM/RBD/RNP motifs) protein	X				2	0.167
miR167	AT3G52190	phosphate transporter traffic facilitator1	X				2	0.167
miR168	AT3G58030	MUSE1	X				1	0.833
miR398	AT1G75270	DHAR2, dehydroascorbate reductase 2	X				2	0.167
miR398	AT2G43900	Endonuclease/exonuclease/phosphatase	X				2	0.167
miR408	AT2G47900	TLP3, TUBBY LIKE PROTEIN 3	X				2	0.167

Table 2.2. Additional TRUEE targets not in the VAT set from only analysing root-specific degradome libraries. The Library % Cut-off threshold meet for each HE target is indicated by 'X'. Bolded genes indicate HE targets which were found to possess T-plots comparable to those in the VAT set. *Maximum Category* (*Max Cat*) indicates whether the highest T-plot Category found across degradome libraries is Cat 1 or Cat2 and *Cat S* is *Category Score*.

2.2.7 Defining the Arabidopsis miRNA targetome

The majority of the literature on *Arabidopsis* MTIs corresponds to the 29 miRNA and tasiRNA families whose targets compose the VAT set. However, this is only a small subset of Arabidopsis miRNAs, as there are 428 annotated miRNAs composing 231 families in Arabidopsis as reported in miRBase v22 (Kozomara et al., 2019). Therefore, to gain a better understanding of the scope of MTIs in Arabidopsis, TRUEE was applied to this complete set of putative Arabidopsis miRNAs (Table S2). The analysis was performed on 34 *Arabidopsis* degradome libraries, which appeared the limit to which WPMIAS could process the 400 miRNAs. The initial analysis was performed at a *Library % Cut-off* of 10% to assist the identification of more subtle MTIs (henceforth, low stringency). The collection of HE targets identified by this analysis is defined as the “miRNA targetome”.

2.2.8 The number of HE targets per miRNA family strongly correlates with miRNA conservation

Given the large numbers of miRNAs, they were first sorted into groups based on conservation (Table 2.3). These conservation-based groups were; (1) miRNAs that have only been identified in *Arabidopsis thaliana* (132 families; referred to as ‘*A. thaliana-specific*’); (2) miRNAs conserved in at least one other species of the *Brassicaceae* (53 families; referred to as ‘*Brassicaceae-specific*’); these included many miRNAs that have only been found in *Arabidopsis thaliana* and *Arabidopsis lyrata*; (3) miRNAs conserved across multiple clades of land plants (27 families; referred to as ‘conserved’), as defined in Axtell and Meyers (2018). Conserved miRNAs were further grouped into *Conserved-guide* (27 families) and *Conserved-passenger* (19 families) as there is evidence that the miRNA passenger strand (miRNA*) also have regulatory roles (Liu et al., 2017).

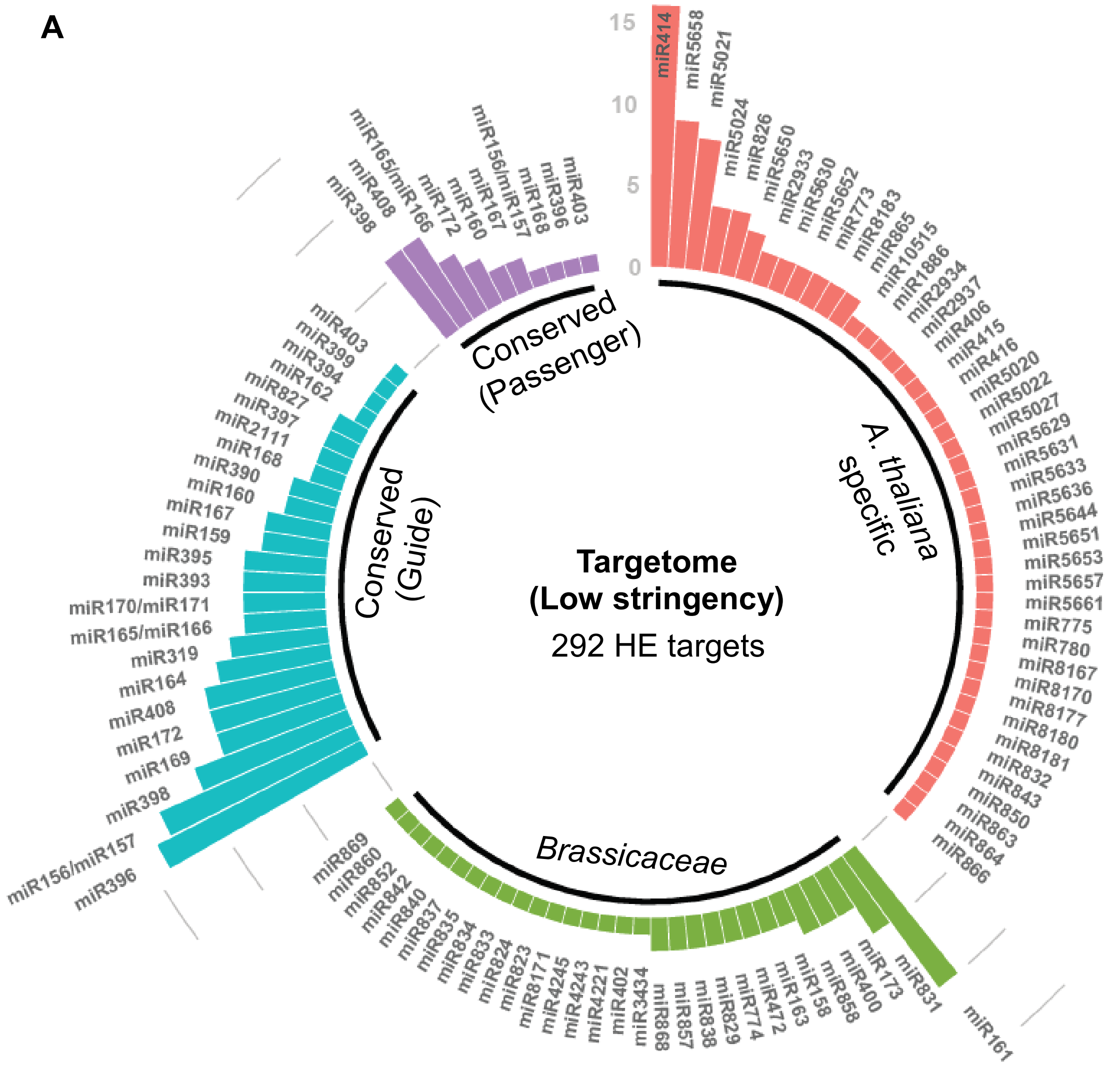
In total, 3478 targets were predicted for the 428 *Arabidopsis* miRNAs by psRNATarget (Table 2.3). Of these, TRUEE identified 292 as HE targets at a low stringency *Library % Cut-off* of 10% (Table 2.3). Therefore, the number of HE targets is at least an order of magnitude lower than the number of predicted targets. The *Conserved-guide* miRNA grouping had the greatest number of HE targets (41%), followed by the *A. thaliana-specific* (30%), *Brassicaceae-specific* (20%), and *Conserved-passenger* (9%) families. Therefore, HE targets of the *Conserved-guide* miRNA group contributes the most to the Arabidopsis targetome, despite this grouping having far fewer miRNA families than the *Brassicaceae-specific* or *A. thaliana-specific* groupings (Table 2.3).

Finally, TRUEE only identified HE targets in 108 of the 231 Arabidopsis miRNA families (Table 2.3). Whereas only 33% of *A. thaliana-specific* families had HE targets, the majority of families in the *Brassicaceae-specific* (30/53; 57%), *Conserved-passenger* (10/19; 53%) and *Conserved-guide* (24/27; 89%) groupings, had HE targets. Therefore, as the conservation of a miRNA family increased, the likelihood it had a HE target increased.

Upon analysing the distribution of HE targets by individual miRNA families, it was found that most *Conserved-guide* families had multiple HE targets (Figure 2.5A). In contrast, most *A. thaliana-specific* and *Brassicaceae-specific* families only had single HE targets, although a few of these families had many HE targets. The *Cat Scores* of the HE targets were determined for each conservation groupings (Figure 2.6A). It was found that the *Cat Scores* for HE targets from the *Conserved-guide* families were the most evenly distributed, ranging from 0.2 to 4.3. By contrast, the number of HE targets for *A. thaliana-specific* and *Brassicaceae-specific* families plateaued around a *Cat Score* of 0.75, and both had relatively few HE targets with a *Cat Score* > 1 (Figure 2.6A). In particular, *Conserved-passenger* families had the fewest HE targets with a *Cat Score* ≥ 0.5, where none exceeded 0.7 (Figure 2.6A; Table S3). Therefore, most of the HE targets with high *Cat Scores* correspond to targets from the *Conserved-guide* families.

miRNA Group	**miRNA[a] families**	**Predicted[b] targets**	**HE targets[c]**		**miRNA families with[d] HE targets**	
			Low	**High**	**Low**	**High**
Conserved-guide	27	493	120	82	24	20
Conserved-passenger	19	478	27	6	10	4
Brassicaceae-specific	53	983	57	19	30	10
A. thaliana-specific	132	1907	88	29	44	14
Total	**231**	**3478**	**292**	**136**	**108**	**48**

Table 2.3. The low and high stringency miRNA targetome of Arabidopsis. [a] the total number of miRNA family entries for *A. thaliana* on miRBase v22 (Kozomara et al., 2019). [b] the number of predicted targets based on default settings of psRNATarget (Dai et al., 2018). [c] the total number of HE targets identified using high and low stringency parameters in TRUEE. [d] the number of miRNA families with HE targets using high and low stringency parameters in TRUEE.


A
15
10
5
0
Targetome
(Low stringency)
292 HE targets
A. thaliana specific
Brassicaceae
Conserved (Guide)
Conserved (Passenger)
miR414
miR5658
miR5021
miR5024
miR826
miR5650
miR2933
miR5630
miR5652
miR773
miR8183
miR865
miR10515
miR1886
miR2934
miR2937
miR406
miR415
miR416
miR5020
miR5022
miR5027
miR5629
miR5631
miR5633
miR5636
miR5644
miR5651
miR5653
miR5657
miR5661
miR775
miR780
miR8167
miR8170
miR8177
miR8180
miR8181
miR832
miR843
miR850
miR863
miR864
miR866
miR161
miR831
miR173
miR400
miR858
miR158
miR163
miR472
miR774
miR829
miR838
miR857
miR868
miR3434
miR402
miR4221
miR4243
miR4245
miR8171
miR823
miR824
miR833
miR834
miR835
miR837
miR840
miR842
miR852
miR860
miR869
miR396
miR156/miR157
miR398
miR169
miR172
miR408
miR164
miR319
miR165/miR166
miR170/miR171
miR393
miR395
miR159
miR167
miR160
miR390
miR168
miR2111
miR397
miR827
miR162
miR394
miR399
miR403
miR398
miR408
miR165/miR166
miR172
miR160
miR167
miR156/miR157
miR168
miR396
miR403

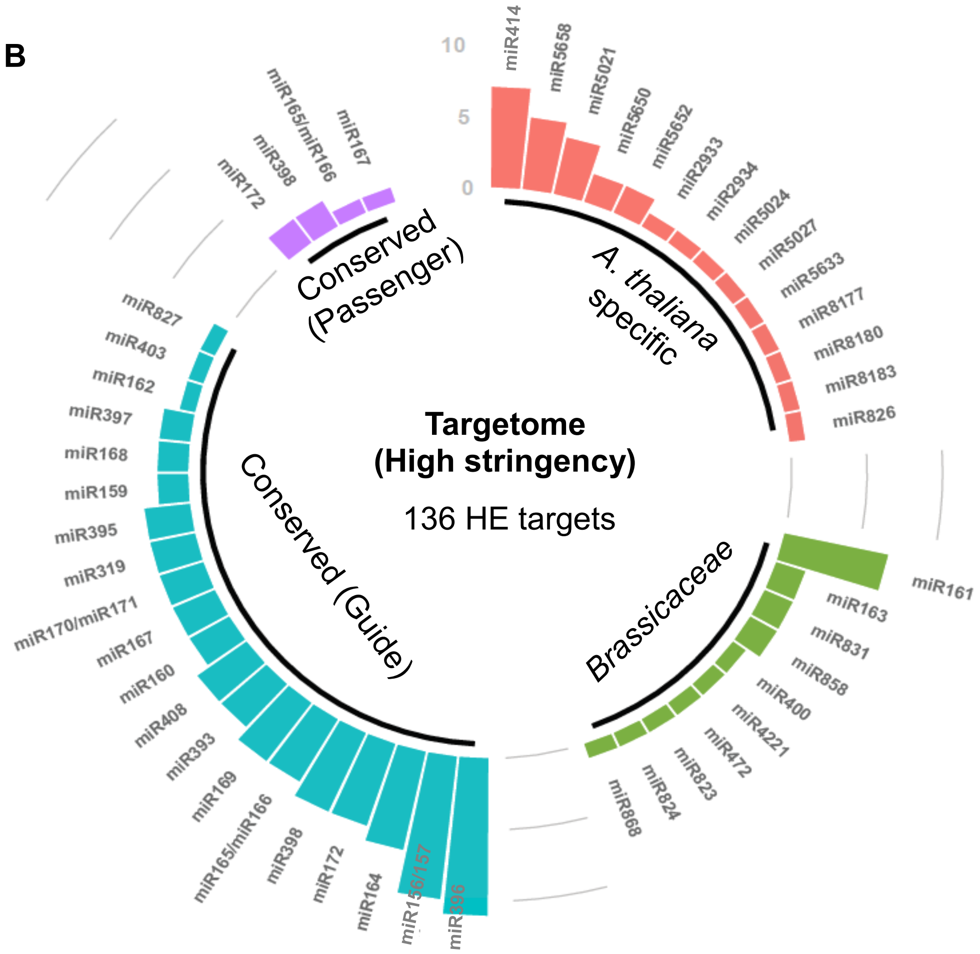


Figure 2.5. The Arabidopsis miRNA targetome. HE targets identified for all Arabidopsis miRNA families by conservation group at; (A) low stringency; (B) high stringency. Families are grouped by conservation so that pink indicates *A. thaliana-specific* miRNA families, green indicates *Brassicaceae-specific* miRNA families, blue indicates *Conserved-guide* miRNA families, and purples indicates *Conserved-passenger* miRNA families. Each bar represents the number of HE targets per miRNA family when analysed by TRUEE.

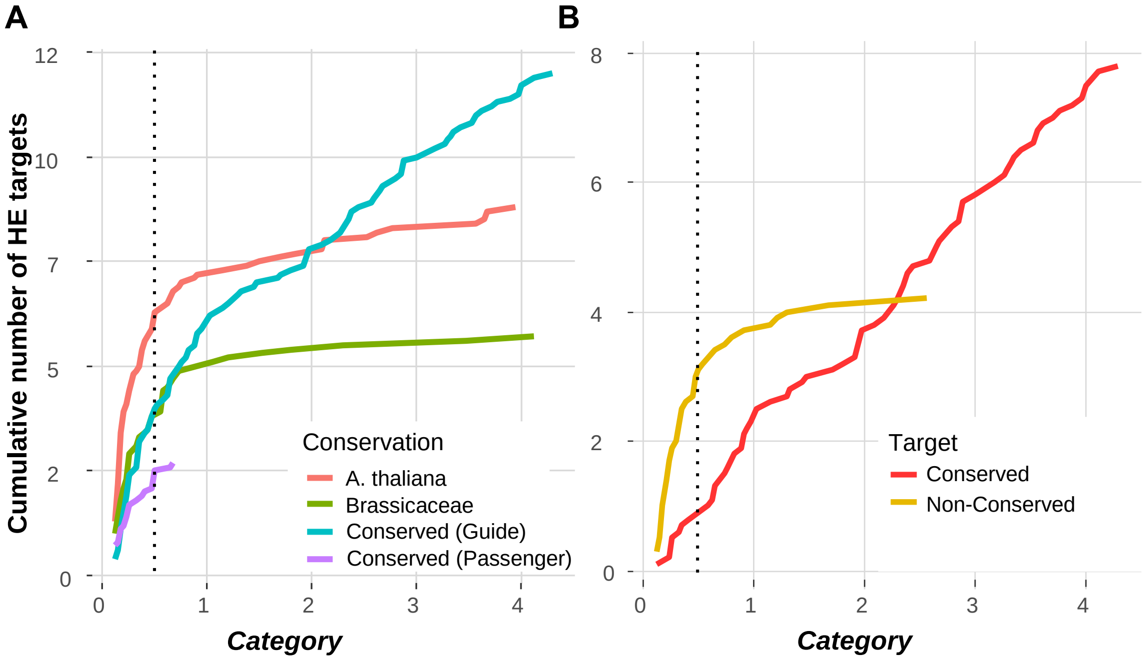


Figure 2.6. The distribution of HE target *Cat Scores* that relate to conservation. (A) The cumulative number of HE targets against *Cat Score* of the different miRNA conservation group. The dotted line indicates a *Cat Score* cut-off of 0.5. (B) The cumulative number of HE targets against *Cat Score* for conserved and non-conserved targets of the *Conserved-guide* miRNA families.

2.2.9 Most HE targets with the highest *Cat Score*s correspond to previously characterised MTIs

Next, the HE targets of *Conserved-guide* miRNA families were classified as either belonging to a conserved target family, or corresponding to being a non-conserved target (Table S4). Most of the HE targets (86%) from conserved target families had a *Cat Score* ≥ 0.5 (Figure 2.6B; Table S4). Alternatively, most non-conserved HE targets (77%) had a *Cat Score* < 0.5 (Figure 2.6B). For non-conserved targets, the highest *Cat Score* was 2.6, whereas many conserved targets exceeded this value, with the highest *Cat Score* being 4.3.

Of the conserved HE targets which had *Cat Scores* ≥ 0.5, all but two were part of the VAT (Table S1), indicating the vast majority of these MTIs have been previous characterised. Interestingly, the only two HE targets not part of the VAT were both homologues of characterised targets; a *NAC* homologue (AT3G12977; miR164) and an *SBP-DOMAIN* homologue (AT5G50670; miR156). For non-conserved targets, the top two HE targets with the highest *Cat Score*s, *RELATED TO AP2 12* (*RAP2.12;* AT1G53910) and *CRY2-INTERACTING BHLH4* (*CIB4*; AT1G10120), were also part of the VAT.

This was also true for the *Brassicaceae-specific* miRNA targets where 15 of the 19 of the HE targets with a *Cat Score* ≥ 0.5 were previously reported as miRNA targets in the literature (either part of the VAT, or otherwise) or were related to these targets (e.g. miR161:*PPR/TPR* family; miR163:*SAMT* family) (Table 2.4., Table S5). Furthermore, the *Brassicaceae-specific* HE targets with the highest *Cat Score*s also corresponded to the most highly studied MTIs, such as the miR161:*PPR/TPR* module and miR824:*AGL16* module (Howell et al., 2007; Kutter et al., 2007; Szaker et al., 2019). By contrast, only four of the 38 *Brassicaceae-specific* HE targets with a *Cat Score* < 0.5 were part of the VAT. Together, these results show that, for the *Conserved-guide* and *Brassicaceae-specific* miRNA groupings, most HE targets with the highest *Cat Score*s are well characterised miRNA targets, or are related to these targets. This argues that the scope of functional MTIs in Arabidopsis has largely been identified.

miRNA	Target ID	*Cat Score*	Previously characterised	Gene Symbol	Target Description
miR161	AT5G41170	4.118	Yes[a]		Pentatricopeptide repeat (PPR-like) superfamily protein
miR824	AT3G57230	3.471	Yes[a]	AGL16	AGAMOUS-like 16
miR823	AT1G69770	2.294	Yes[a]	CMT3	CHROMOMETHYLASE 3
miR161	AT1G06580	1.794	Yes[a]		Pentatricopeptide repeat (PPR) superfamily protein
miR472	AT5G43740	1.529	Yes[b]		Disease resistance protein (CC-NBS-LRR class) family
miR163	AT1G66700	1.206	Yes[a]	PXMT1	S-adenosyl-L-methionine-dependent methyltransferase
miR4221	AT1G20500	1.059	No		AMP-dependent synthetase and ligase family protein
miR161	AT1G64583	1.059	Yes[a]		Tetratricopeptide repeat (TPR)-like superfamily protein
miR400	AT1G62720	0.735	Yes[c]		Pentatricopeptide repeat (PPR-like) superfamily protein
miR831	AT3G56020	0.735	No		Ribosomal protein L41 family
miR868	AT1G18270	0.676	No		ketose-bisphosphate aldolase class-II family protein
miR831	AT3G08520	0.676	No		Ribosomal protein L41 family
miR858	AT2G47460	0.618	Yes[d]	MYB12	myb domain protein 12
miR161	AT1G62910	0.588	Yes[a]		Pentatricopeptide repeat (PPR) superfamily protein
miR161	AT1G62914	0.588	Yes[a]		pentatricopeptide (PPR) repeat-containing protein
miR161	AT1G62930	0.588	Yes[a]		Tetratricopeptide repeat (TPR)-like superfamily protein
miR161	AT1G63130	0.588	Yes[a]		Tetratricopeptide repeat (TPR)-like superfamily protein
miR163	AT3G44860	0.588	Yes[a]	FAMT	farnesoic acid carboxyl-O-methyltransferase
miR858	AT4G26930	0.559	MYB12 related	MYB97	myb domain protein 97

[a] Part of or related to targets in the VAT set

[b] Boccara et al., 2014

[c] Park et al., 2014

[d] Sharma et al., 2016

Table 2.4. HE targets of *Brassicaceae-specific* miRNA families with a *Cat Score* ≥0.5. List of HE targets with [a] indicating that the target is part of, or related to genes in the VAT set. [b, c, d] Indicate genes that are not in the VAT set but are supported in literature to have genetic and molecular evidence as miRNA targets.

2.2.10 Many HE targets of *A. thaliana-specific* miRNAs are diverse genes with trinucleotide repeats

By contrast, most of the HE targets for the *A. thaliana-specific* families have not been previously described, and none were present in the VAT set. Of the 29 HE targets with *Cat Scores* ≥ 0.5, 16 were targets of three miRNAs, miR414, miR5021, and miR5658, with some of these HE targets having very strong *Cat Scores* (Table 2.5). Curiously, all three miRNAs are mainly composed of repeated trinucleotide sequences which was also characteristic of their binding sites in their HE targets. Additionally, the HE targets of miR414, miR5021, and miR5658 HE targets did not appear to be related in identity, but rather diverse mRNA targets containing these trinucleotide repeats.

miRNA	Rep. miRNA	Target ID	Cat Score	Gene Symbol	Target description
miR414	Yes	AT5G55580	3.941		Mitochondrial transcription termination factor
miR5021	Yes	AT2G40520	3.676		Nucleotidyltransferase family protein
miR5021	Yes	AT5G24670	3.676		Cytidine/deoxycytidylate deaminase
miR5021	Yes	AT1G03190	3.647	UVH6	RAD3-like DNA-binding helicase protein
miR5021	Yes	AT3G23890	3.559	TOPII	topoisomerase II
miR414	Yes	AT5G40340	2.765		Tudor/PWWP/MBT superfamily protein
miR8177		AT1G15710	2.618		prephenate dehydrogenase family protein
miR5652		AT1G62670	2.529	RPF2	rna processing factor 2[a]
miR414	Yes	AT3G11810	2.118		(1 of 2) PTHR33133:SF7 - F26K24.10
miR414	Yes	AT5G55300	2.118	TOP1ALPHA	DNA topoisomerase I alpha
miR414	Yes	AT1G16150	2.088	WAKL4	wall associated kinase-like 4
miR414	Yes	AT1G60220	1.853	ULP1D	UB-like protease 1D
miR5658	Yes	AT1G73710	1.706		Pentatricopeptide repeat (PPR) superfamily
miR5658	Yes	AT4G11600	1.5	GPX6	glutathione peroxidase 6
miR5658	Yes	AT5G56860	1.382	GNC	GATA type zinc finger transcription factor
miR5633		AT2G35670	1.147	FIS2	VEFS-Box of polycomb protein
miR5652		AT5G16640	0.912		Pentatricopeptide repeat (PPR) superfamily[a]
miR5027		AT1G07610	0.882	MT1C	metallothionein 1C
miR2933		AT4G32390	0.765		Nucleotide-sugar transporter family protein
miR5658	Yes	AT2G32310	0.735		CCT motif family protein
miR2934		AT5G03650	0.676	SBE2.2	starch branching enzyme 2.2
miR8183		AT5G04220	0.676	SYTC	Calcium-dependent lipid-binding family
miR414	Yes	AT5G64830	0.676		PCD 2 C-terminal domain-containing protein
miR5658	Yes	AT4G20070	0.618	AAH	allantoate amidohydrolase
miR5650		AT5G03240	0.618	UBQ3	polyubiquitin 3
miR826		AT1G09730	0.5	ASP1	Arabidopsis sumo protease 1
miR5024		AT3G57290	0.5	EIF3E	Eukaryotic translation initiation factor 3
miR8180		AT4G29350	0.5	PFN2	Profilin 2
miR5650		AT5G20620	0.5	UBQ4	ubiquitin 4

Table 2.5. HE targets of *A. thaliana*-specific miRNA families with a *Cat Score* ≥ 0.5. List of the HE targets, and indication of whether it is regulated by a miRNA with trinucleotide repeats (Rep. miRNA), *Cat Score*, and gene annotation. None of these targets are in the VAT set.

2.2.11 A high stringency Arabidopsis miRNA targetome

Given the above analyses have shown the majority of MTIs with strong experimental evidence correspond to HE targets with *Library % Cut-off* of 10% and a *Cat Score* cut-off of ≥ 0.5, imposing these cut-offs appears justified in terms of capturing MTIs with the highest evidence in a bid to define a high stringency Arabidopsis targetome. Using these parameters, only 136 HE targets are identified, with the *Conserved-guide* HE targets now making up a majority of targets (60%), followed by *A. thaliana-specific* (21%), *Brassicaceae-specific* (14%), and *Conserved-passenger* (5%) families (Figure 2.5B). In this high stringency targetome, the number of miRNA families with HE targets dropped to only 48 of the 231 miRNA families (21%), with the *A. thaliana-specific* (14/132; 11%), *Brassicaceae-specific* (10/53; 19%) and *Conserved-passenger* (4/19; 21%) groupings now all having a minority of miRNA families with HE targets. This reduction stems mainly from the exclusion of single HE target-miRNA interactions being filtered from this high stringency Arabidopsis targetome (Figure 2.5B). By contrast, a majority of *Conserved-guide* families still had HE targets (20/27; 74%). Hence, TRUEE is filtering out a set of targets that is in line with the long-standing notion that most functional MTIs are conserved (Axtell, 2008), rather than the possibility of promiscuous targeting of many mRNA via a large and diverse miRNome (Brodersen and Voinnet, 2009).

2.3 Discussion

A central question of plant miRNA biology is the identification of functionally important (physiologically relevant) MTIs. Here, TRUEE has been developed to filter and rank MTIs based on experimental evidence. This was then applied to Arabidopsis as a proof-of-concept to define an accurate estimation of the number of functional MTIs in a plant, termed the "miRNA targetome". Although non-exhaustive, the approach suggests Arabidopsis would have no more than 300 functionally MTIs, and likely, considerably fewer. In the context of this paper, functionally important refers to an MTI that if altered, would alter a plant trait (i.e. have a physiological impact).

2.3.1 TRUEE; a simple approach to rank MTIs independently of miRNA-target complementarity

We aimed to develop a simple bioinformatic approach based on currently available and widely utilized online tools. Firstly, psRNATarget is the most widely used and cited plant miRNA target prediction tool that has been recently updated (Dai et al., 2018). It is a highly accessible, user-friendly webserver, and is compatible with WPMIAS (Fei et al., 2021). WPMIAS is also a highly accessible, user-friendly webserver and currently is the simplest tool to analyse multiple degradome libraries.

Unlike previous miRNA target prediction tools that are based on miRNA-target complementarity, the scoring schema of TRUEE is derived solely from degradome data. It is based on the strength and frequency of a target's T-plots across multiple degradome libraries, from which the *Cat Score* can be derived, a metric that directly relates to extent of miRNA-mediated cleavage. Like WPMIAS (Fei et al., 2021), *Target Categories* 3 and 4 were not considered strong enough evidence for miRNA-mediated cleavage (so are essentially given a weighted score of 0). This approach is justified in that using only *Target Categories* 1 and 2 was sufficient to identify the vast majority of the VAT set (Figure 2.3). *Target Category* 1 was given an arbitrary weighted value 5-fold greater than *Target Category* 2 plots given the much greater confidence that these signals are derived from miRNA-mediated regulation. This is illustrated in Figure 2.4, where it is unclear whether the *Target Category* 2 signals for *POR C* and *CYCLOPHILIN 38* is derived from miRNA-mediated cleavage or other degradation mechanisms.

Finally, if TRUEE is compared to data from the most recently published tool, TarHunter (Ma et al., 2018), it appears TRUEE is identifying less false positives. Using TarHunter in the ortho_mode (protein and nucleotide sequence at the target site is conserved) and most stringent number of mismatches, TarHunter identifies 59 targets for the conserved set of miRNAs in Arabidopsis (http://www.biosequencing.cn/TarHunter/ath.html). Of these, 17 (29%) are not present in the VAT. Therefore, even at the highest stringency of TarHunter, it appears that TRUEE is identifying proportionally fewer false positives.

2.3.2 Limitations of TRUEE

Firstly, given the presence of a degradome signal requires both the presence of the miRNA and transcription of the target mRNA, TRUEE will preferentially detect MTIs that are widespread, and potentially miss those MTIs that have a narrow temporal and spatial occurrence. Both the canonical nutrient dependent miR399:*PHO2* and miR395:*SULTR2* MTIs had low *Cat Score*s (0.265; Table S4), as the majority of the degradome analyses have likely not been performed when conditions exist for these MTIs. To potentially offset this, selection of particular degradome libraries (conditions or tissues), may help identify these narrow MTIs, as was demonstrated for the root MTIs. The current code (published on Open Science Framework) is customisable, so that the analysis of any subset(s) of degradome libraries is possible. As most degradome experiments are only a snapshot of miRNA-mediated activity at one particular developmental stage or growth condition, obviously the larger the number of degradome libraries analysed, the more comprehensive a picture will be of the miRNA targetome.

Secondly, TRUEE will not detect targets which are regulated solely by translational repression. However, this may be inconsequential, as nearly all canonical targets were identified using TRUEE, validating the use of this approach to detect the vast majority of miRNA targets. This is consistent with the observation that canonical targets that are known to undergo translational repression, are also cleaved by the miRNA (for review see Yu et al., 2017), implying there is no strong evidence that miRNA targets are solely regulated by translational repression.

Thirdly, using the libraries provided by psRNATarget and WPMIAS, no non-coding RNA can be identified (eg. miR390:*TAS3*). Therefore, miRNA:long non-coding RNA interactions cannot be considered in this study.

2.3.3 The functional miRNA targetome of Arabidopsis

Currently, the functional scope of the plant miRNome remains contentious. As many studies claim that most miRNAs in a plant are lineage-specific (Cui et al., 2017), and that many of these miRNA–target interactions are evolutionarily fluid (Smith et al., 2015), these notions align with the hypothesis that there are likely 100s of functional miRNAs and 1000s of MTIs. However, other researchers are more cautious, and question the validity of many of these species-specific miRNAs that have been annotated on databases such as miRBase (Axtell and Meyers, 2018) or argue that most non-conserved miRNAs are likely to be evolutionary transient with no functional targets (Axtell, 2008; Cuperus et al., 2011). In this study, by determining how many functional MTIs there are in a plant and the proportion of these that correspond to non-conserved miRNAs, we aimed to add weight to which hypothesis is more likely.

Our findings support the notion that only several 100 MTIs of functional importance are present in a plant (Li et al., 2014; Taylor et al., 2014). Although previously proposed, the value of reiterating this notion has merit in that many current studies assume there are 1000s of MTIs of functional importance as predicted by bioinformatics (Lindow and Krogh, 2005; Lindow et al., 2007; Dai et al., 2018; Bülow et al., 2012, Kozomara et al., 2019, Fei et al., 2020). Moreover, without the filters imposed by TRUEE, studies based on degradome data also claim 1000s of targets [e.g. WPMIAS reports >10 000 MTIs in *Oryza sativa* from an analysis of 738 miRNAs (Fei et al., 2020)]. Our findings align with the view of Axtell and Meyers (2018), in that prediction of 1000s of targets, followed by Gene Ontology or KEGG Ontology analysis to infer miRNA function is problematic (Eldem et al., 2012; Yaish et al., 2015; Yawichai et al., 2019; Tiwari et al., 2020; Xu et al., 2020), and likely has little relevance to miRNA function *in planta*. We advocate that using an approach such as TRUEE will enable to rapid identification of which genes are being strongly regulated by miRNA, and therefore, what genetic targets would be best to modify in the bid to improving desired plant traits.

Our analyses support the idea that the majority of functional MTIs have already been identified in Arabidopsis. In the analysis of 34 Arabidopsis degradome libraries in WPMIAS (Fei et al., 2020), the known conserved canonical miRNA targets had the highest-ranking *Cat Score*s, indicating this metric was able to filter out and identify strong MTIs that have clear functional roles (Table S4). By contrast, there were very few uncharacterised MTIs that had a high *Cat Score*. This extended to the *Brassicaceae-specific* MTIs, where the highest *Cat Score*s were largely limited to previously documented MTIs, such as the well-studied miR824:*AGL16* and miR161:*PPR* modules (Howell et al., 2007; Kutter et al., 2007; Szaker et al., 2019).

It could be argued that only a subset of sRNAs were investigated, as the complex miRNome includes miRNA isoforms that arise through altered processing or modifications and that are predicted to confer altered specificity, and these were not included in the analysis. To investigate this possibility, we analysed the passenger strands (miRNA*s) of conserved guide miRNAs, as currently this class of alternative miRNA isoforms have the strongest evidence implicating them in functional MTIs (Zhang et al., 2011; Manavella et al., 2013; Du et al., 2017; Liu et al., 2017). However, only a few HE targets were identified for this *Conserved-passenger* grouping and all had low *Cat Score*s (< 0.7). Moreover, previous reported functional miRNA*-target interactions, such as miR393* (Zhang et al., 2011) were not detected in the analysis. Again, it is possible that these classes of sRNAs have highly specific temporal and/or spatial expression, and so their MTIs are missed due to the absence of the corresponding degradome libraries, as TRUEE will be biased towards MTIs that are widespread. Nevertheless, despite the regulatory potential of the miRNA*s, none of their MTIs have *Cat Score*s characteristics of the known physiologically important MTIs.

For the majority of Arabidopsis miRNA entries in miRBase, TRUEE either failed to identify a HE target (72% - *Brassicaceae-specific* and 89% for the *A. thaliana-specific* groupings) or had a single-target with a low *Cat Score*. This is consistent with the observation that most low confidence miRNA entries on miRBase corresponded to poorly expressed, evolutionarily young miRNAs that lack a functional target gene (Cuperus et al., 2011), and the annotation of many of these being *bona fide* miRNAs has been questioned (Taylor et al., 2017; Axtell and Meyers, 2018). It is consistent with the hypothesis of the existence within the plant cell of a large pool of diverse, evolutionarily young, and weakly expressed miRNAs from which new MTIs of functional significance may arise (Rajagopalan et al., 2006; Fahlgren et al., 2007; Axtell et al., 2007; Axtell, 2008; Cuperus et al., 2011). However, it has been hypothesised this is rare and most young miRNAs remain targetless and undergo neutral drift until their sequences are no longer recognisable by DCL for processing (Axtell, 2008; Cuperus et al., 2011). Again, it may be argued that many young MTIs will not be identified by TRUEE because they have a narrow spatial and temporal expression. However, that any young MTI can be detected, such as miR824:*AGL16*, which are localised in stomatal complexes, suggests otherwise (Kutter et al., 2007).

Finally, the highest ranking HE targets of the *A. thaliana-specific* miRNAs, predominantly consisted of targets of three unrelated miRNAs that have trinucleotide repeats, miR414, miR5021 and miR5658. For each miRNA, their targets consisted of diverse genes with the common feature of trinucleotide repeats at their potential binding site. Trinucleotide repeat expansions are known to cause multiple human genetic diseases such as Huntington’s disease and has been reported to cause sensitivity to high temperatures in the *A. thaliana* accession *Bur-0* (Bates et al., 2015; Tabib et al., 2016). Therefore, these *A. thaliana-specific* miRNAs may have a specialised role in silencing potentially deleterious genes with trinucleotide repeat expansions. However, these claims will need to be tested with experimental analyses.

2.3.4 Conclusions

TRUEE represents an approach to rank miRNA-targets independently of complementarity, circumventing the limitation of that approach that has been a central feature of bioinformatic target prediction programs. We envision the approach can be applied to other species, once sufficient degradome analyses have been conducted. It will enable fast ranking of targets, and therefore, which genes to modify in regard to the plant traits that miRNAs control.

2.4 Experimental Procedure

2.4.1 Bioinformatics workflow

The parameters of TRUEE were developed via benchmarking the retrieval of the VAT set. The VAT set was assembled via systematically and manually reviewing the literature, requiring two independent lines of evidence from commonly used experimental approaches. The literature supporting the formation of the VAT set is found in Table S1.

Mature miRNA sequences were retrieved from miRBase v22 (Kozomara et al., 2019). Where multiple isomiRs were found, the isomiR with the highest abundance found on a plant next-generation sequencing database (https://mpss.danforthcenter.org) was used (Nakano et al., 2020). The most conserved tasiARF sequence as reported by Allen et al. (2005) was used in the analysis. For the Arabidopsis “miRNA targetome”, all 428 available mature miRNA sequences which includes isomiRs, were retrieved from miRBase v22 (Kozomara et al., 2019; note that tasiARFs were not analysed as they are not on miRbase).

Sequences were used as input into psRNATarget v2, 2017 scoring schema (Dai et al., 2018). Default settings were used for analysis other than the expectation score which was decreased to 3 for all sRNAs except miR167, miR398 and miR408. An expectation score of 5 was used for these miRNAs as their targets from the VAT set exceeds an expectation score of 3.

The resulting predicted targets were then analysed using WPMIAS (Fei et al., 2020). WPMIAS settings were; (1) Analysis type - Analysis > Advanced II > Use psRNATarget predicted results directly; (2) Plant species - *Arabidopsis thaliana*; cDNA libraries - Transcript, JGI genomic project, Phytozome 11, 167 TAIR10 (from psRNAtarget); (3) Offset from spliced position (nt) - 0 (default), or 1 for miR162, miR396 and miR398 which can only be identified using an offset of 1 (Yamasaki et al., 2007; Debernardi et al., 2012; Shao et al., 2015); (4) Mismatches allowed for mapping degradome reads to references: - 0 (default).

Degradome data retrieved from WPMIAS was then used as input and analysed using TRUEE to identify HE and LE targets as described in Figure 2.1. TRUEE was developed using an in-house R script. Analysed data from WPMIAS and R script for TRUEE is accessible on the Open Science Framework page for this project https://osf.io/k7rcs/. Target Categories as defined in WPMIAS were used in this study (Fei et al., 2020).

2.4.2 Data visualization

Multiple sequence alignments (MSAs) were performed using Multiple Alignment using Fast Fourier Transform (MAFFT) (Katoh and Standley, 2013), and the resulting alignment visualised using Jalview (Waterhouse et al., 2009). T-plots of miRNA targets were adapted from WPMIAS

(Fei et al., 2020). Figures determining the optimal expectation score (Figure 2.2), identifying the HE targets by TRUEE (Figure 2.3), and the Arabidopsis targetome (Figure 2.5) were generated using R package, ggplot2. Code and design for Figure 2.5 was by Holtz Yan and can be found at https://www.r-graph-gallery.com/297-circular-barplot-with-groups.html. All graphs were generated using the R package, ggplot2.

References

Addo-Quaye, C., Eshoo, T. W., Bartel, D. P. and Axtell, M. J. (2008) Endogenous siRNA and miRNA targets identified by sequencing of the Arabidopsis degradome. *Current Biology*, 18, 758–762. https://doi.org/10.1016/j.cub.2008.04.042

Allen, E., Xie, Z., Gustafson, A. M., Sung, G. H., Spatafora, J. W. and Carrington, J. C. (2004) Evolution of microRNA genes by inverted duplication of target gene sequences *in Arabidopsis thaliana*. *Nature Genetics*, 36, 1282–1290. https://doi.org/10.1038/ng1478

Allen, E., Xie, Z., Gustafson, A. M. and Carrington, J. C. (2005) microRNA-directed phasing during trans-acting siRNA biogenesis in plants. *Cell*, 121, 207–221. https://doi.org/10.1016/j.cell.2005.04.004

Allen, R. S., Li, J., Stahle, M. I., Dubroué, A., Gubler, F. and Millar, A. A. (2007) Genetic analysis reveals functional redundancy and the major target genes of the Arabidopsis miR159 family. *Proc Natl Acad Sci USA*, 104, 16371–16376. https://doi.org/10.1073/pnas.0707653104

Axtell, M. J., Snyder, J. A. and Bartel, D. P. (2007) Common functions for diverse small RNAs of land plants. *Plant Cell*, 19, 1750–1769. https://doi.org/10.1105/tpc.107.05170

Axtell, M. J. (2008) Evolution of microRNAs and their targets: Are all microRNAs biologically relevant? Biochimica et Biophysica Acta, 1779, 725–734. https://doi.org/10.1016/j.bbagrm.2008.02.007

Axtell, M. J. and Meyers, B. C. (2018) Revisiting criteria for plant microRNA annotation in the era of big data. *Plant Cell*, 30, 272–284. https://doi.org/10.1105/tpc.17.00851

Bates, G. P., Dorsey, R., Gusella, J. F., Hayden, M. R., Kay, C., Leavitt, B. R., Nance, M., Ross, C. A., Scahill, R. I., Wetzel, R., Wild, E. J. and Tabrizi, S. J. (2015) Huntington disease. *Nature Reviews Disease Primers*, 1, 15005. https://doi.org/10.1038/nrdp.2015.5

Boccara, M., Sarazin, A., Thiébeauld, O., Jay, F., Voinnet, O., Navarro, L. and Colot, V. (2014) The Arabidopsis miR472-RDR6 Silencing Pathway Modulates PAMP- and Effector-Triggered Immunity through the Post-transcriptional Control of Disease Resistance Genes. *PLoS Pathogens*, 10, e1003883. https://doi.org/10.1371/journal.ppat.1003883

Bonnet, E., He, Y., Billiau, K. and Van de Peer, Y. (2010) TAPIR, a web server for the prediction of plant microRNA targets, including target mimics. *Bioinformatics*, 26, 1566–1568. https://doi.org/10.1093/bioinformatics/btq233

Brodersen, P. and Voinnet, O. (2009) Revisiting the principles of microRNA target recognition and mode of action. *Nature Reviews Molecular Cell Biology*, 10, 141–148. https://doi.org/10.1038/nrm2619

Brousse, C., Liu, Q., Beauclair, L., Deremetz, A., Axtell, M. J. and Bouché, N. (2014) A non-canonical plant microRNA target site. *Nucleic Acids Research*, 42, 5270–5279. https://doi.org/10.1093/nar/gku157

Bülow, L., Bolívar, J. C., Ruhe, J., Brill, Y. and Hehl, R. (2012) 'MicroRNA Targets', a new AthaMap web-tool for genome-wide identification of miRNA targets in *Arabidopsis thaliana*. *BioData Mining*, 5, 7. https://doi.org/10.1186/1756-0381-5-7

Chorostecki, U., Crosa, V. A., Lodeyro, A. F., Bologna, N. G., Martin, A. P., Carrillo, N., Schommer, C. and Palatnik, J. F. (2012) Identification of new microRNA-regulated genes by conserved targeting in plant species. *Nucleic Acids Research*, 40, 8893–8904. https://doi.org/10.1093/nar/gks625

Cui, J., You, C., and Chen, X. (2017) The evolution of microRNAs in plants. *Current opinion in plant biology*, 35, 61–67. https://doi.org/10.1016/j.pbi.2016.11.006

Cuperus, J. T., Fahlgren, N. and Carrington, J. C. (2011) Evolution and functional diversification of MIRNA genes. *Plant Cell*, 23, 431–442. https://doi.org/10.1105/tpc.110.082784

Dai, X. and Zhao, P. X. (2011) PsRNATarget: A plant small RNA target analysis server. *Nucleic Acids Research*, 39(SUPPL. 2), 155–159. https://doi.org/10.1093/nar/gkr319

Dai, X., Zhuang, Z. and Zhao, P. X. (2018) PsRNATarget: A plant small RNA target analysis server (2017 release). *Nucleic Acids Research*, 46(W1), W49–W54. https://doi.org/10.1093/nar/gky316

Debernardi, J. M., Rodriguez, R. E., Mecchia, M. A. and Palatnik, J. F. (2012) Functional specialization of the plant miR396 regulatory network through distinct microRNA-target interactions. *PLoS Genetics*, 8. e1002419 https://doi.org/10.1371/journal.pgen.1002419

Du, Q., Zhao, M., Gao, W., Sun, S. and Li, W.-X. (2017) microRNA/microRNA* complementarity is important for the regulation pattern of NFYA5 by miR169 under dehydration shock in Arabidopsis. *Plant Journal*, 91, 22–33. https://doi.org/https://doi.org/10.1111/tpj.13540

Eldem, V., Çelikkol Akçay, U., Ozhuner, E., Bakır, Y., Uranbey, S. and Unver, T. (2012) Genome-wide identification of miRNAs responsive to drought in peach (Prunus persica) by high-throughput deep sequencing. *PLoS One*, 7, e50298. https://doi.org/10.1371/journal.pone.0050298

Fahlgren, N., Howell, M. D., Kasschau, K. D., Chapman, E. J., Sullivan, C. M., Cumbie, J. S., Givan, S. A., Law, T. F., Grant, S. R., Dangl, J. L. and Carrington, J. C. (2007) High-throughput sequencing of Arabidopsis microRNAs: evidence for frequent birth and death of MIRNA genes. *PLoS One*, 2, e219. https://doi.org/10.1371/journal.pone.0000219

Fei, Y., Mao, Y., Shen, C., Wang, R., Zhang, H. and Huang, J. (2020) WPMIAS: Whole-degradome-based Plant MicroRNA-Target Interaction Analysis Server. *Bioinformatics* (Oxford, England), 36, 1937-1939. https://doi.org/10.1093/bioinformatics/btz820

Folkes, L., Moxon, S., Woolfenden, H. C., Stocks, M. B., Szittya, G., Dalmay, T. and Moulton, V. (2012) PAREsnip: a tool for rapid genome-wide discovery of small RNA/target interactions evidenced through degradome sequencing. *Nucleic Acids Research*, 40, e103. https://doi.org/10.1093/nar/gks277

German, M.A., Pillay, M., Jeong, D.-H., Hetawal, A., Luo, S., Janardhanan, P., Kannan, V., Rymarquis, L. a, Nobuta, K., German, R., De Paoli, E., Lu, C., Schroth, G., Meyers, B. C. and Green, P. J. (2008) Global identification of microRNA-target RNA pairs by parallel analysis of RNA ends. *Nature Biotechnology*, 26, 941–946. https://doi.org/10.1038/nbt1417

Howell, M. D., Fahlgren, N., Chapman, E. J., Cumbie, J. S., Sullivan, C. M., Givan, S. A., Kasschau, K. D. and Carrington, J. C. (2007) Genome-wide analysis of the RNA-DEPENDENT RNA POLYMERASE6/DICER-LIKE4 pathway in Arabidopsis reveals dependency on miRNA- and tasiRNA-directed targeting. *Plant Cell*, 19, 926–942. https://doi.org/10.1105/tpc.107.050062

Jeong, D.-H., Schmidt, S. A., Rymarquis, L. A., Park, S., Ganssmann, M., German, M. A., Accerbi, M., Zhai, J., Fahlgren, N., Fox, S. E., Garvin, D. F., Mockler, T. C., Carrington, J. C., Meyers, B. C. and Green, P. J. (2013) Parallel analysis of RNA ends enhances global investigation of microRNAs and target RNAs of *Brachypodium distachyon*. *Genome Biology*, 14, R145–R145. https://doi.org/10.1186/gb-2013-14-12-r145

Katoh, K. and Standley, D. M. (2013) MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution*, 30, 772–780. https://doi.org/10.1093/molbev/mst010

Kozomara, A., Birgaoanu, M. and Griffiths-Jones, S. (2019) miRBase: from microRNA sequences to function. *Nucleic Acids Research*, 47(D1), D155–D162. https://doi.org/10.1093/nar/gky1141

Kutter, C., Schöb, H., Stadler, M., Meins, F. and Si-Ammour, A. (2007) MicroRNA-mediated regulation of stomatal development in Arabidopsis. *Plant Cell*, 19, 2417–2429. https://doi.org/10.1105/tpc.107.050377

Li, J., Reichel, M., Li, Y. and Millar, A. A. (2014) The functional scope of plant microRNA-mediated silencing. *Trends in Plant Science*, 19, 750–756. https://doi.org/10.1016/j.tplants.2014.08.006

Lindow, M. and Krogh, A. (2005) Computational evidence for hundreds of non-conserved plant microRNAs. *BMC Genomics*, 6, 119. https://doi.org/10.1186/1471-2164-6-119

Lindow, M., Jacobsen, A., Nygaard, S., Mang, Y. and Krogh, A. (2007) Intragenomic Matching Reveals a Huge Potential for miRNA-Mediated Regulation in Plants. *PLOS Computational Biology*, 3, e238. https://doi.org/10.1371/journal.pcbi.0030238

Liu, Q., Wang, F. and Axtell, M. J. (2014) Analysis of Complementarity Requirements for Plant MicroRNA Targeting Using a *Nicotiana benthamiana* Quantitative Transient Assay. *Plant Cell*, 26, 741–753. https://doi.org/10.1105/tpc.113.120972

Liu, W. W., Meng, J., Cui, J. and Luan, Y. S. (2017) Characterization and function of microRNA*s in plants. *Frontiers in Plant Science*, 8, 1–7. https://doi.org/10.3389/fpls.2017.02200

Ma, X., Liu, C., Gu, L., Mo, B., Cao, X. and Chen, X. (2018) TarHunter, a tool for predicting conserved microRNA targets and target mimics in plants. *Bioinformatics*, 34, 1574–1576. https://doi.org/10.1093/bioinformatics/btx797

Mallory, A. C., Reinhart, B. J., Jones-Rhoades, M. W., Tang, G., Zamore, P. D., Barton, M. K. and Bartel, D. P. (2004) MicroRNA control of PHABULOSA in leaf development: Importance of pairing to the microRNA 5′ region. *EMBO Journal*, 23, 3356–3364. https://doi.org/10.1038/sj.emboj.7600340

Manavella, P. A., Koenig, D., Rubio-Somoza, I., Burbano, H. A., Becker, C. and Weigel, D. (2013) Tissue-specific silencing of Arabidopsis SU(VAR)3-9 HOMOLOG8 by miR171a. *Plant physiology*, 161, 805–812. https://doi.org/10.1104/pp.112.207068

Nakano, M., McCormick, K., Demirci, C., Demirci, F., Gurazada, S. G. R., Ramachandruni, D., Dusia, A., Rothhaupt, J. A. and Meyers, B. C. (2020) Next-Generation Sequence Databases: RNA and Genomic Informatics Resources for Plants. *Plant Physiology*, 182, 136–146. https://doi.org/10.1104/pp.19.00957

Park, Y. J., Lee, H. J., Kwak, K. J., Lee, K., Hong, S. W. and Kang, H. (2014) MicroRNA400-Guided Cleavage of Pentatricopeptide Repeat Protein mRNAs Renders Arabidopsis thaliana More Susceptible to Pathogenic Bacteria and Fungi. *Plant and Cell Physiology*, 55, 1660–1668. https://doi.org/10.1093/pcp/pcu096

Rajagopalan, R., Vaucheret, H., Trejo, J. and Bartel, D. P. (2006) A diverse and evolutionarily fluid set of microRNAs in Arabidopsis thaliana. *Genes and Development*, 20, 3407–3425. https://doi.org/10.1101/gad.1476406

Schwab, R., Palatnik, J. F., Riester, M., Schommer, C., Schmid, M. and Weigel, D. (2005) Specific effects of microRNAs on the plant transcriptome. *Developmental Cell*, 8, 517–527. https://doi.org/10.1016/j.devcel.2005.01.018

Shao, F., Qiu, D. and Lu, S. (2015) Comparative analysis of the Dicer-like gene family reveals loss of miR162 target site in SmDCL1 from *Salvia miltiorrhiza*. *Scientific Reports* 5, 9891. https://doi.org/10.1038/srep09891

Sharma, D., Tiwari, M., Pandey, A., Bhatia, C., Sharma, A. and Trivedi, P. K. (2016) MicroRNA858 Is a Potential Regulator of Phenylpropanoid Pathway and Plant Development. *Plant Physiology*, 171, 944–959. https://doi.org/10.1104/pp.15.01831

Smith, L. M., Burbano, H. A., Wang, X., Fitz, J., Wang, G., Ural-Blimke, Y., and Weigel, D. (2015). Rapid divergence and high diversity of miRNAs and miRNA targets in the *Camelineae*. *Plant Journal,* 81, 597–610. https://doi.org/10.1111/tpj.12754

Srivastava, P. K., Moturu, T. R., Pandey, P., Baldwin, I. T. and Pandey, S. P. (2014) A comparison of performance of plant miRNA target prediction tools and the characterization of features for genome-wide target prediction. *BMC Genomics*, 15, 348. https://doi.org/10.1186/1471-2164-15-348

Sun, Y. H., Lu, S., Shi, R. and Chiang, V. L. (2011) Computational prediction of plant miRNA targets. Methods in molecular biology (Clifton, N.J.), 744, 175–186. https://doi.org/10.1007/978-1-61779-123-9_12

Szaker, H. M., Darkó, É., Medzihradszky, A., Janda, T., Liu, H. C., Charng, Y. Y. and Csorba, T. (2019) miR824/AGAMOUS-LIKE16 Module Integrates Recurring Environmental Heat Stress Changes to Fine-Tune Poststress Development. *Frontiers in Plant Science*, 10, 1–19. https://doi.org/10.3389/fpls.2019.01454

Tabib, A., Vishwanathan, S., Seleznev, A., McKeown, P. C., Downing, T., Dent, C., Sanchez-Bermejo, E., Colling, L., Spillane, C. and Balasubramanian, S. (2016) A Polynucleotide Repeat Expansion Causing Temperature-Sensitivity Persists in Wild Irish Accessions of *Arabidopsis thaliana*. *Frontiers in Plant Science*, 7, 1311. https://www.frontiersin.org/article/10.3389/fpls.2016.01311

Taylor, R. S., Tarver, J. E., Hiscock, S. J. and Donoghue, P. C. J. (2014) Evolutionary history of plant microRNAs. *Trends in Plant Science*, 19, 175–182. https://doi.org/10.1016/j.tplants.2013.11.008

Taylor, R. S., Tarver, J. E., Foroozani, A. and Donoghue, P. C. J. (2017) Insights and Perspectives MicroRNA annotation of plant genomes. Do it right or not at all. *BioEssays*, 1600113, 1–6. https://doi.org/10.1002/bies.201600113

Thody, J., Moulton, V. and Mohorianu, I. (2020) PAREameters: a tool for computational inference of plant miRNA–mRNA targeting rules using small RNA and degradome sequencing data. *Nucleic Acids Research*, 48, 2258–2270. https://doi.org/10.1093/nar/gkz1234

Tiwari, J. K., Buckseth, T., Zinta, R., Saraswati, A., Singh, R. K., Rawat, S. and Chakrabarti, S. K. (2020) Genome-wide identification and characterization of microRNAs by small RNA sequencing for low nitrogen stress in potato. *PLoS ONE*, 15, e0233076. https://doi.org/10.1371/journal.pone.0233076

Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. and Barton, G. J. (2009) Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, 25, 1189–1191. https://doi.org/10.1093/bioinformatics/btp033

Xu, X., Chen, X., Chen, Y., Zhang, Q., Su, L., Chen, X., Chen, Y., Zhang, Z., Lin, Y. and Lai, Z. (2020) Genome-wide identification of miRNAs and their targets during early somatic embryogenesis in *Dimocarpus longan* Lour. *Scientific Reports*, 10, 4626. https://doi.org/10.1038/s41598-020-60946-y

Yamasaki, H., Abdel-Ghany, S. E., Cohu, C. M., Kobayashi, Y., Shikanai, T. and Pilon, M. (2007) Regulation of copper homeostasis by micro-RNA in Arabidopsis. *Journal of Biological Chemistry*, 282, 16369–16378. https://doi.org/10.1074/jbc.M700138200

Yaish, M. W., Sunkar, R., Zheng, Y., Ji, B., Al-Yahyai, R. and Farooq, S. A. (2015) A genome-wide identification of the miRNAome in response to salinity stress in date palm (*Phoenix dactylifera* L.). *Front Plant Science*, 6, 946. https://www.frontiersin.org/article/10.3389/fpls.2015.00946

Yawichai, A., Kalapanulak, S., Thammarongtham, C. and Saithong, T. (2019) Genome-Wide Identification of Putative MicroRNAs in Cassava (Manihot esculenta Crantz) and Their Functional Landscape in Cellular Regulation. *BioMed Research International*, 2019846. https://doi.org/10.1155/2019/2019846

Yu, Y., Jia, T. and Chen, X. (2017) The 'how' and 'where' of plant microRNAs. *New Phytologist*, 216, 1002–1017. https://doi.org/10.1111/nph.14834

Zhang, X., Zhao, H., Gao, S., Wang, W. C., Katiyar-Agarwal, S., Huang, H. D., Raikhel, N. and Jin, H. (2011) Arabidopsis Argonaute 2 regulates innate immunity via miRNA393(∗)-mediated silencing of a Golgi-localized SNARE gene, MEMB12. *Molecular Cell*, 42, 356–366. https://doi.org/10.1016/j.molcel.2011.04.010

Zheng, Y., Li, Y. F., Sunkar, R. and Zhang, W. (2012) SeqTar: An effective method for identifying microRNA guided cleavage sites from degradome of polyadenylated transcripts in plants. *Nucleic Acids Research*, 40, e28. https://doi.org/10.1093/nar/gkr1092

Zheng, Z., Reichel, M., Deveson, I., Wong, G., Li, J. and Millar, A. A. (2017) Target RNA Secondary Structure Is a Major Determinant of miR159 Efficacy. *Plant Physiology*, 174, 1764–1778. https://doi.org/10.1104/pp.16.01898

Chapter 3

Conserved plant miRNAs: identifying their targets across the plant kingdom and the factors impacting their specificity

Abbreviations

AGO – Argonaute

AP2 – APETELA2-LIKE

ARF – AUXIN RESPONSE FACTOR

APS – ATP-SULFURYLASE

CDS – coding sequence

COPT – copper transporter

COX – CYTOCHROME C OXIDASE

CSD1 – COPPER SUPEROXIDE DISMUTASE

DCL – DICER-LIKE

FLU – FLUORESCENT IN BLUE LIGHT

GRF3 – GROWTH-REGULATING FACTOR 3

HAM – HAIRY MERISTEM

HD-ZIPIII – CLASS III HOMEODOMAIN LEUCINE ZIPPER

HE – High Evidence

IAR3 – IAA-ALANINE RESISTANT 3

IPS1 – INDUCED BY PHOSPHATE STARVATION1

LAC – LACCASE

LB – Luria Broth

LE – Low Evidence

MAFFT – Multiple Alignment using Fast Fourier Transform

miRISC – miRNA-induced silencing complex

miRNA – microRNAs

MSAs – Multiple Sequence Alignments

MTIs – miRNA-Target Interactions

NF-YA – NUCLEAR TRANSCRIPTION FACTOR Y SUBUNIT ALPHA

NLA – NITROGEN LIMITATION ADAPTATION

nts – nucleotides

PANTHER ID – Protein ANalysis THrough Evolutionary Relationships ID

PHO2 – PHOSPHATE2

PHT5 – PLASMA-MEMBRANE-LOCALIZED PHOSPHATE TRANSPORTER 5

phyloP – phylogenetic *P*-values

Pi – phosphate

rPHAST – Phylogenetic Analysis with Space/Time Models

SL1 – stem-loop 1

SL2 – stem-loop 2

SOD – SUPEROXIDE DISMUTASE

SPX – SYG1/PHO81/XPR1

SULTR – SULFATE TRANSPORTER

T-plots – target plots

TAS3 – TRANS-ACTING SHORT INTERFERING RNA 3

TCP – TEOSINTE BRANCHED1, CYCLOIDEA, AND PROLIFERATING CELL NUCLEAR ANTIGEN BINDING FACTOR

TRUEE – Targets Ranked Using Experimental Evidence

WPMIAS – Whole-Degradome-based Plant MicroRNA-Target Interaction Analysis Server

Abstract

In plants, high complementarity between microRNAs (miRNAs) and their target genes is a prerequisite for a miRNA-target interaction (MTI). However, evidence suggests there are factors beyond complementarity that impacts the strength of the MTI. To explore this, the bioinformatic pipeline TRUEE (Targets Ranked Using Experimental Evidence) was applied to a set of conserved miRNAs to identity their high evidence (HE) targets across species. For each conserved miRNA family, HE targets mostly consisted of homologues from one conserved target gene family (termed the "primary family"). If an additional HE target family(s) was identified ("secondary family"), it was often functionally related to the primary family, suggesting plant miRNAs preferentially regulate functionally related genes. Multiple sequence alignments of homologues of primary families found highly conserved sequences flanking their miRNA-binding sites. These conserved flanking sequences were enriched in homologues found in the HE target set across species, suggesting they facilitate miRNA-mediated regulation. Curiously, a subset of these flanking sequences was predicted to form conserved RNA secondary structures that preferentially involved base-pairing with the miRNA-binding sites, counterintuitive to the notion that functional miRNA-binding sites need to be unstructured and highly accessible for strong miRNA-mediated regulation. Finally, functional testing of the conserved flanking sequences of the miR160 target, *AUXIN RESPONSE FACTOR 10* (*ARF10*), found that mutations within these flanking sequences resulted in attenuated *ARF10* silencing by miR160. Together, these findings suggest that many of these ancient miRNA-target relationships have developed regulatory complexities beyond complementarity that define them as strongly regulated, functional target genes of miRNAs.

3.1 Introduction

The most commonly reported plant microRNAs (miRNAs) in the literature correspond to a set of several dozen miRNAs families that are highly conserved across land plants (Axtell and Meyers, 2018). From nearly two decades of study, it is clear each of these conserved miRNA families have a single highly conserved family of target genes (Schwab et al., 2005; reviewed in Jones-Rhoades, 2012; reviewed in Tang & Chu, 2017). Underpinning the conservation of these miRNA-target interactions (MTIs) is that they are largely involved in core biological processes in plants, such as fundamental developmental processes (e.g. miR156, miR160, miR165/166, miR172) (Mallory et al., 2004; Mallory et al., 2005; Palatnik et al., 2007; Wang et al., 2009), and abiotic and biotic stress responses (e.g. miR395, miR397, miR398) (Morel et al., 2002; Sunkar et al., 2006; Abdel-Ghany & Pilon, 2008; Kawashima et al., 2009). As the identity of these conserved target families are predominantly regulatory genes such as transcription factors and F-box proteins, these conserved miRNAs have the potential to regulate entire gene expression programs (reviewed in Jones-Rhoades et al., 2006). Highlighting their importance, perturbation of many of these MTIs leads to mutant phenotypes with pleiotropic defects (Todesco et al., 2010). Consistently, the previous Chapter found that these conserved MTIs are the highest-ranking Targets Ranked Using Experimental Evidence (TRUEE) targets, and therefore have the highest evidence as miRNA targets.

It has long been known that plant MTIs require a high degree of complementarity (Rhoades et al., 2002; Schwab et al., 2005; Liu et al., 2014). Based on this, many bioinformatic prediction tools to identify miRNA targets generally assume that mRNAs with a high complementarity miRNA-binding site equates to a genuine target gene. Although this approach has successfully identified most conserved targets (Rhoades et al., 2002; Jones-Rhoades & Bartel, 2004), many conserved targets are found to have a lower degree of complementarity than many predicted targets for which no experimental validation exists (Table 1.1 of Chapter 1). Therefore, complementarity alone is insufficient in ranking which genes are subject to physiologically relevant miRNA-regulation, implying factors other than complementarity are involved. Given the evolutionary age of these MTIs, it is feasible that such additional regulatory factors could arise.

Currently however, there is very little evidence to support the existence of such factors. Several studies have investigated the possibility that miRNA-binding sites are present in highly accessible regions of the target transcripts. Firstly, it was bioinformatically shown that across multiple species, AU rich synonymous codons were enriched in the 96 nucleotides (nts) flanking upstream and downstream of the miRNA-binding sites (Gu et al., 2012). This also correlated with a greater miRNA-binding site accessibility and suggests a reduction in RNA secondary structures is being selected for. However, this analysis was performed on psRNAtarget predicted targets and so the data may be compromised by the preponderance of false positives. Furthermore, a study on the RNA secondary structure of the *Arabidopsis thaliana* (henceforth, Arabidopsis) transcriptome found the 21 nt miRNA-binding site to be less structured compared to the 50 nt sequences immediately flanking upstream and downstream of this region (Li et al., 2012).

However, as an *in vitro* study, conclusions drawn from this study must be taken in the context that it was conducted in the absence of RNA-Binding Proteins and other cellular influences. Indeed, apposing this study, a recent *in vivo* study found miRNA-binding sites to be highly structured, with their unfolding being the limiting factor of cleavage efficiency directed by a miRNA-induced silencing complex (miRISC) (Yang et al., 2020). Here, only the two nts immediately downstream of the miRNA-binding site were required to be single stranded for efficient cleavage (Yang et al., 2020). Supporting this notion of a highly structured miRNA-binding site, was the discovery of highly conserved RNA secondary structures associated with the miR159-binding site of two *GAMYB* genes in Arabidopsis (*MYB33* and *MYB65*), and that were functionally demonstrated to promote miR159-mediated silencing in Arabidopsis (Zheng et al., 2017). Given this occurs independently of AU content or predicted miRNA-binding site accessibility, it highlights our lack of understanding of features promoting miRNA-mediated silencing (Zheng et al., 2017).

The presence of the RNA secondary structures associated with the miR159-binding site of the MYB genes supports the previous hypothesis that the long evolutionary history of these ancient MTIs may have enabled additional regulatory mechanisms beyond miRNA-target complementarity to arise (Li et al., 2014). From sequence analysis, it is apparent that these conserved MTIs are fixed across species, with multiple members of a specific miRNA family having high complementarity to multiple members of a specific target family (Li et al., 2014; Axtell & Meyers, 2018). Given the dominance of these highly conserved target families, it was previously suggested that they could be considered as the primary target(s) of these conserved miRNA families (Li et al., 2014). It was further hypothesized, that as an active miRISC is all that is needed to execute silencing, the acquisition of any additional targets would need to be compatible with the parameters of the primary miRNA-target relationship, and therefore this would likely limit the promiscuity of functional miRNA-targeting (Li et al., 2014).

In this Chapter, I aim to explore these hypotheses by applying TRUEE to determine the identity of High Evidence (HE) targets of the highly conserved miRNAs across diverse plant species. The identified HE targets will then be examined to identify potential conserved features beyond sequence complementarity. Finally, I will functionally test whether these features can contribute to the efficacy to which the target is subject to miRNA-mediate regulation.

This Chapter aims to find:

1) Are the conserved target families the predominant targets of conserved miRNAs across species?
2) How often are additional target families acquired by conserved miRNAs, if at all, and the identities of these targets?
3) The extent that miRNA-target binding site complementarity can be used as an indicator of MTIs corresponding to HE targets.
4) Are there features additional to miRNA-binding site complementarity that are determinants of HE targets?

3.2 Results

3.2.1 HE targets primarily consist of a single gene family for most conserved miRNA

TRUEE was applied to 21 highly conserved miRNAs and a tasiARF across diverse plant species to identify HE and Low Evidence (LE) targets using the parameters described in the Material and Methods. As expected, LE targets outnumbered HE targets for most miRNAs across species (Figure 3.1). The exception was miR169, where LE targets consisted half the total predicted targets. For miRNAs in which the expectation score was increased from 3.0 to 5.0, the number of LE target increased by almost an order of magnitude.

Next, to determine the number of gene families targeted by the conserved miRNAs, the gene family of HE and LE targets were identified using their associated PANTHER ID (Protein ANalysis THrough Evolutionary Relationships) (Mi et al., 2013). Results show that for each miRNA and tasiARF, HE targets across species were predominantly composed of homologues of the same gene family (Figure 3.2). These families were the same as those most often reported in literature to be targets of their corresponding miRNA, and hence, considered here as the primary target family (Jones-Rhoades & Bartel, 2004; Jones-Rhoades, 2012; Sunkar et al., 2012; Chauhan et al., 2017). For some miRNAs, HE targets are almost exclusively made up of homologues of this primary target family; miR160 (94% - *AUXIN RESPONSE FACTOR* (*ARF*)), miR166 (99% - *CLASS III HOMEODOMAIN LEUCINE ZIPPER* (*HD ZIPIII*)), miR170/miR171 (97% - *HAIRY MERISTEM* (*HAM*)), miR172 (95% - *APETELA2-LIKE* (*AP2*)) and tasiARF (98% - *ARF*). This demonstrates that these miRNAs-target relationships are fixed across species.

In contrast, LE targets predominantly consisted of genes from diverse PANTHER IDs. This is indicated in that they largely consisted of gene families which are grouped in the ‘other’ category (PANTHER IDs associated with three or less members) (Figure 3.2). In all but one miRNA (miR397), the ‘other’ category made up over half of the total LE targets. Furthermore, in all cases, the percentage of primary target families in LE targets were smaller than in the HE targets. Strikingly, for some miRNA families, no or very few primary target family members are found to be LE targets; this includes miR162, miR166, miR167, miR168, miR393, miR394, miR398 and miR403. In fact, no primary target family members were identified as LE targets for miR394; only one for miR162; and two for miR167 and miR393. This highlights the prevalence to which these primary target families are subjected to miRNA-mediated regulation.

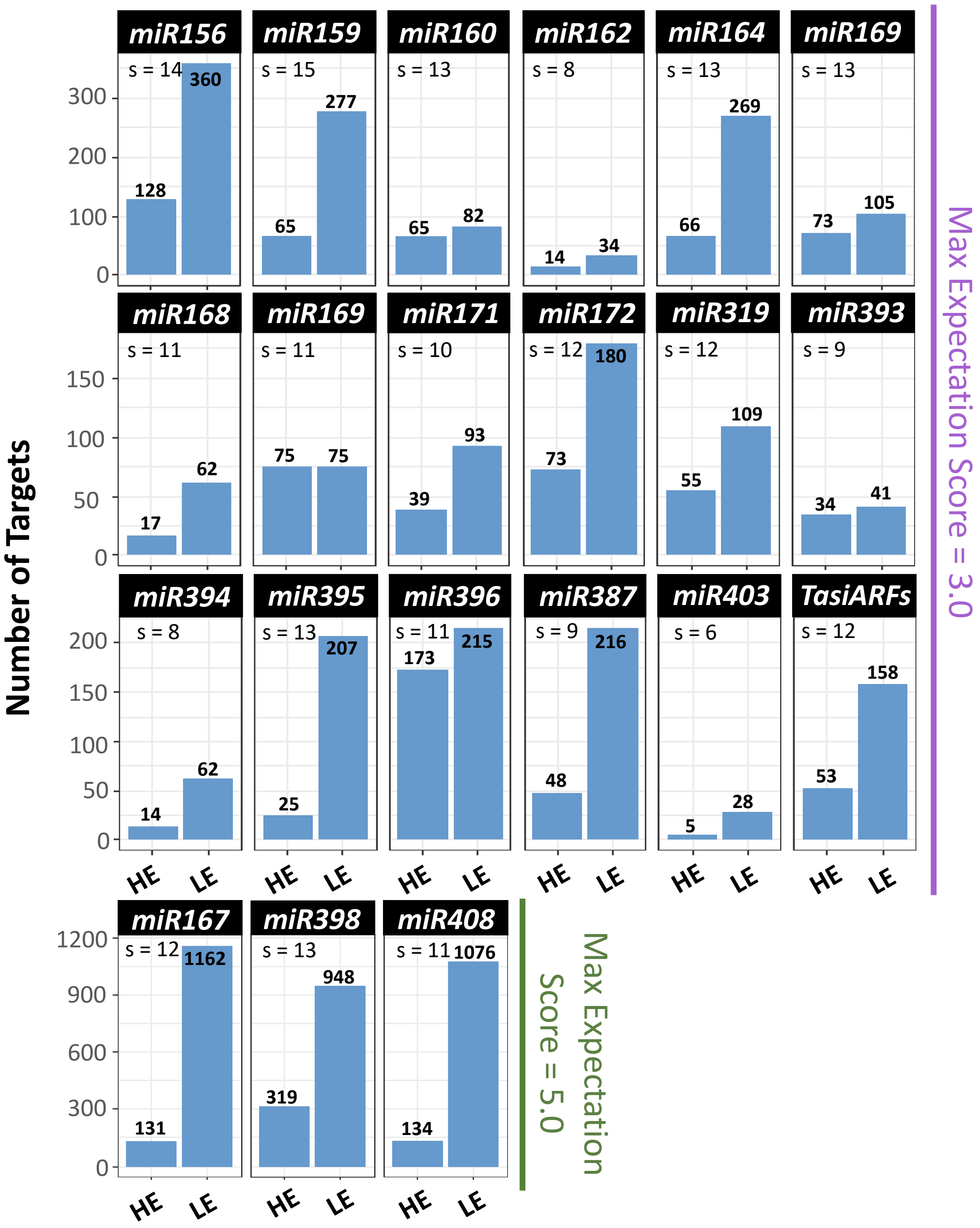
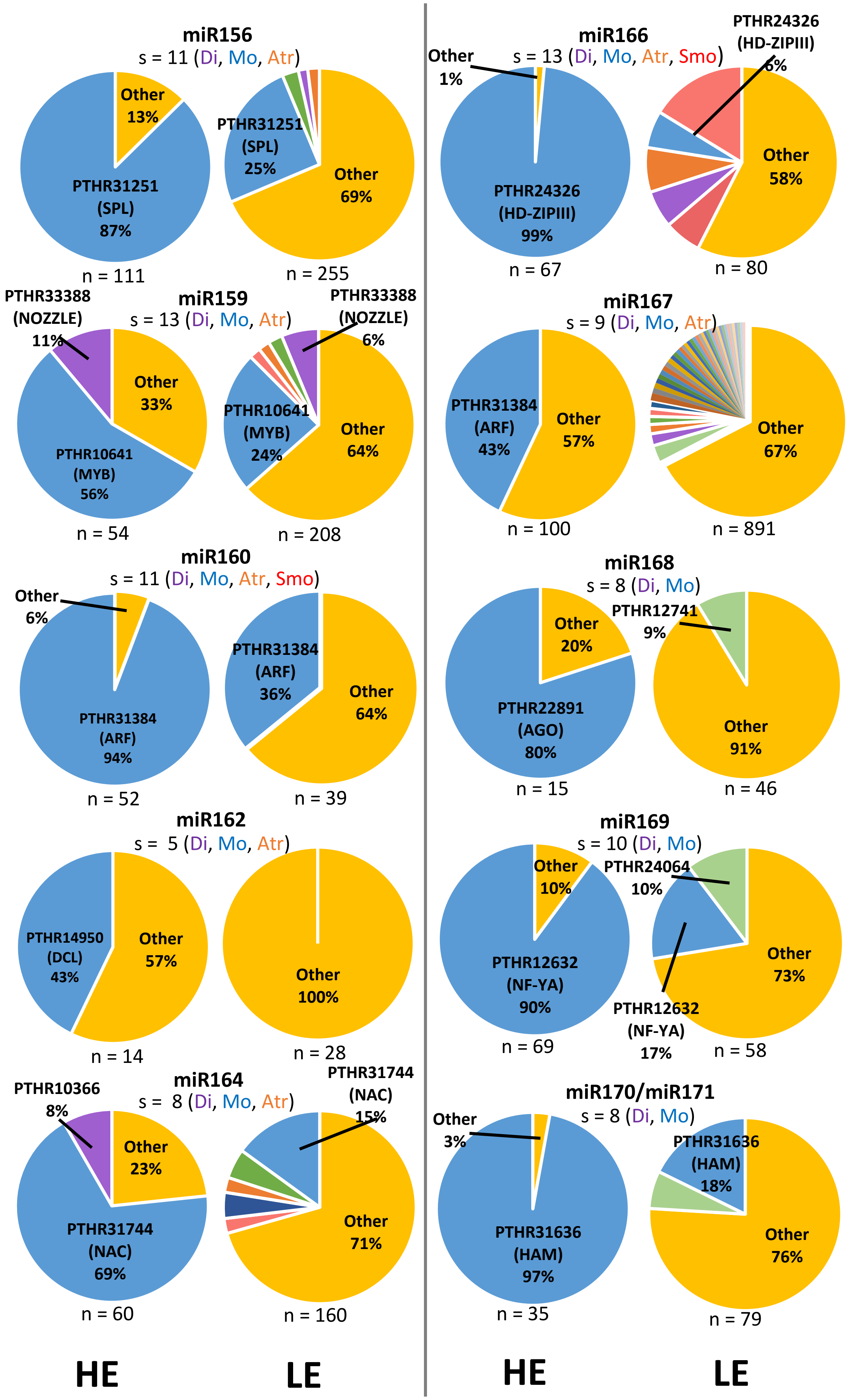
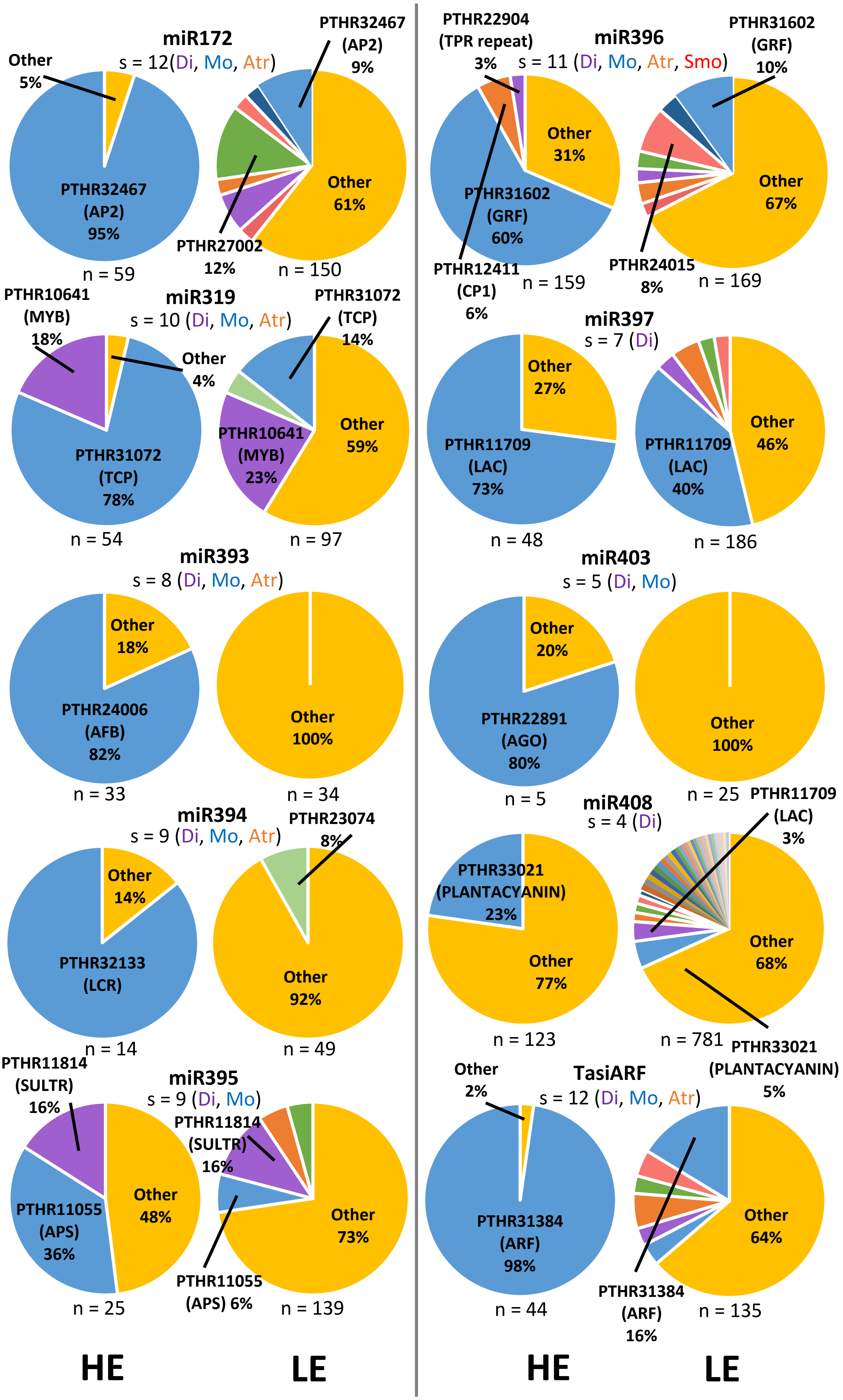


Figure 3.1. HE and LE targets for conserved miRNAs from diverse species. The number of target genes across species with HE and LE targets for 21 conserved miRNAs and TasiARF. Numbers on the top of bars indicate the number of genes found across all species. The expectation score for target prediction with psRNATarget was increased to 5.0 (green) for miR167, miR398 and miR408 as the known targets for Arabidopsis exceed an expectation score of 3.0. ‘s’ indicates the number of species analysed per miRNA. Note the differences in scale.

3.2.2 Few HE targets are found outside the primary target family

For the HE targets, six miRNA families (miR159, miR164, miR319, miR395, miR396 and miR398) regulate additional conserved target families (defined as having four or more conserved HE targets across multiple plant species) (Figure 3.2). These additional HE target families, henceforth called secondary target families, had fewer HE targets compared to the primary target families and were found in a narrower range of species (Table 3.1). For instance, whereas HE targets from all primary target families were found beyond dicotyledonous species, HE targets from the secondary target families were restricted to dicotyledonous species with the exception of miR319:*MYB*. However, miR319 is closely related to miR159 and can both target *MYB* genes, although in Arabidopsis, it was shown that targeting of *MYB* by miR319 is limited, with miR159 being the major regulator (Palatnik et al., 2007). This also appears conserved across species, as the miR319-mediated *MYB* gene regulation detected is much weaker than the corresponding miR159-mediated regulation as indicated by their respective target plots (T-plots) (Figure 3.3). Therefore, the general trend remains, where a conserved miRNA family predominantly regulates one primary target family that is conserved across species, and although acquisition of secondary target families occurs, targeting of these families is less conserved and fewer homologues are regulated.


miR156
s = 11 (Di, Mo, Atr)
Other 13%
PTHR31251 (SPL) 87%
n = 111
PTHR31251 (SPL) 25%
Other 69%
n = 255
miR166
s = 13 (Di, Mo, Atr, Smo)
Other 1%
PTHR24326 (HD-ZIPIII) 99%
n = 67
PTHR24326 (HD-ZIPIII) 6%
Other 58%
n = 80
miR159
s = 13 (Di, Mo, Atr)
PTHR33388 (NOZZLE) 11%
Other 33%
PTHR10641 (MYB) 56%
n = 54
PTHR33388 (NOZZLE) 6%
PTHR10641 (MYB) 24%
Other 64%
n = 208
miR167
s = 9 (Di, Mo, Atr)
PTHR31384 (ARF) 43%
Other 57%
n = 100
Other 67%
n = 891
miR160
s = 11 (Di, Mo, Atr, Smo)
Other 6%
PTHR31384 (ARF) 94%
n = 52
PTHR31384 (ARF) 36%
Other 64%
n = 39
miR168
s = 8 (Di, Mo)
Other 20%
PTHR22891 (AGO) 80%
n = 15
PTHR12741 9%
Other 91%
n = 46
miR162
s = 5 (Di, Mo, Atr)
PTHR14950 (DCL) 43%
Other 57%
n = 14
Other 100%
n = 28
miR169
s = 10 (Di, Mo)
Other 10%
PTHR12632 (NF-YA) 90%
n = 69
PTHR24064 10%
Other 73%
PTHR12632 (NF-YA) 17%
n = 58
miR164
s = 8 (Di, Mo, Atr)
PTHR10366 8%
Other 23%
PTHR31744 (NAC) 69%
n = 60
PTHR31744 (NAC) 15%
Other 71%
n = 160
miR170/miR171
s = 8 (Di, Mo)
Other 3%
PTHR31636 (HAM) 97%
n = 35
PTHR31636 (HAM) 18%
Other 76%
n = 79
HE
LE
HE
LE


miR172
s = 12(Di, Mo, Atr)
Other
5%
PTHR32467
(AP2)
95%
n = 59
PTHR32467
(AP2)
9%
Other
61%
PTHR27002
12%
n = 150
miR396
s = 11 (Di, Mo, Atr, Smo)
PTHR22904
(TPR repeat)
3%
Other
31%
PTHR31602
(GRF)
60%
PTHR12411
(CP1)
6%
n = 159
PTHR31602
(GRF)
10%
Other
67%
PTHR24015
8%
n = 169
miR319
s = 10 (Di, Mo, Atr)
PTHR10641
(MYB)
18%
Other
4%
PTHR31072
(TCP)
78%
n = 54
PTHR31072
(TCP)
14%
PTHR10641
(MYB)
23%
Other
59%
n = 97
miR397
s = 7 (Di)
Other
27%
PTHR11709
(LAC)
73%
n = 48
Other
46%
PTHR11709
(LAC)
40%
n = 186
miR393
s = 8 (Di, Mo, Atr)
Other
18%
PTHR24006
(AFB)
82%
n = 33
Other
100%
n = 34
miR403
s = 5 (Di, Mo)
Other
20%
PTHR22891
(AGO)
80%
n = 5
Other
100%
n = 25
miR394
s = 9 (Di, Mo, Atr)
Other
14%
PTHR32133
(LCR)
n = 14
PTHR23074
8%
Other
92%
n = 49
miR408
s = 4 (Di)
PTHR33021
(PLANTACYANIN)
23%
Other
77%
n = 123
PTHR11709
(LAC)
3%
Other
68%
n = 781
PTHR33021
(PLANTACYANIN)
5%
miR395
s = 9 (Di, Mo)
PTHR11814
(SULTR)
16%
PTHR11055
(APS)
36%
Other
48%
n = 25
PTHR11814
(SULTR)
16%
Other
73%
PTHR11055
(APS) 6%
n = 139
TasiARF
s = 12 (Di, Mo, Atr)
Other
2%
PTHR31384
(ARF)
98%
n = 44
Other
64%
PTHR31384
(ARF)
16%
n = 135
HE
LE
HE
LE

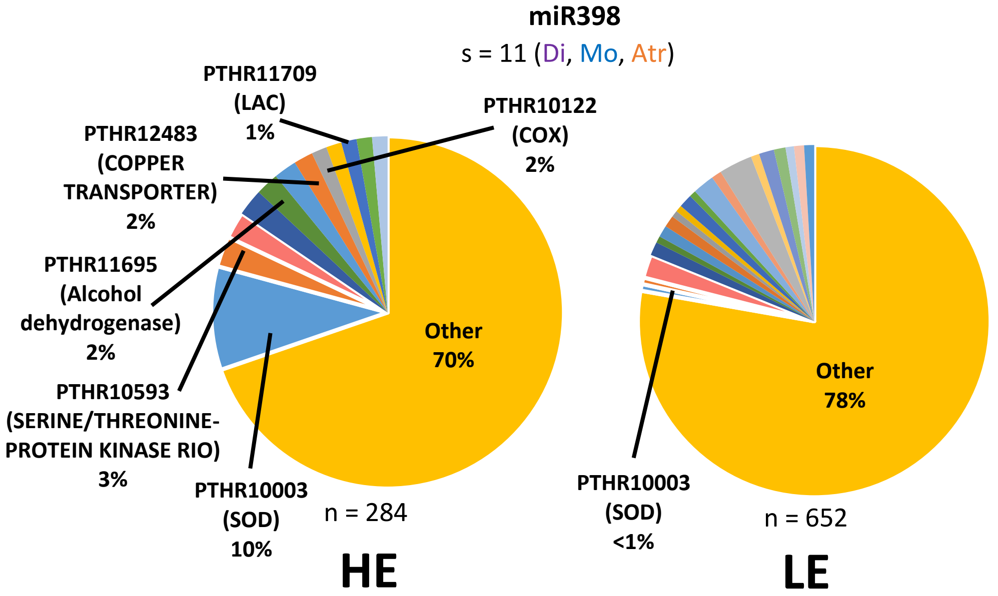


Figure 3.2. Distribution of the gene families of HE and LE targets per miRNA and TasiARF. HE and LE targets were categorised into gene families by their associated PANTHER ID. PANTHER IDs with 3 or less members were grouped into the ‘other’ category. ‘n’ indicates the number of HE or LE targets used in the analysis. ‘s’ indicates the number of species used in the analysis, where ‘Di’ indicates dicots, ‘Mo’ monocots, ‘Atr’ *Amborella trichopoda*, and ‘Smo’ *Selanginella moellendorffii*. The PANTHER ID and gene family name is indicated. Targets with no associated PANTHER ID are not included in the analysis, hence the total number of targets is less than in Figure 3.1.

		Ath	Csi	Gma	Mdm	Mtr	Ppe	Sly	Vvi	Bdi	Hvu	Osa	Zma	Atr
miR159	PTHR10641 (*MYB*)	2	2	6	1	2	1	3	2	1	4	2	2	2
	PTHR33388 (*NOZZLE*)	1		3			1	1						
miR164	**PTHR31744 (*NAC*)**	**5**	**1**	**10**	**4**	**3**	**2**	**3**	**3**	**4**		**3**	**1**	**2**
	PTHR10366 (*NAD DEPENDENT EPIMERASE/DEHYDRATASE*)			1			1	3						
miR319	**PTHR31072 (*TCP*)**	**5**		**10**	**8**	**2**	**2**	**4**	**2**			**4**	**4**	**1**
	PTHR10641 (*MYB*)			6					2			1	1	
miR395	**PTHR11055 (*APS*)**	**2**	**1**		**2**		**1**	**2**	**1**	**1**				
	PTHR11814 (*SULTR*)			3					1					
miR396	**PTHR31602 (*GRF*)**	**6**	**6**	**24**		**9**	**6**	**10**		**9**		**11**	**9**	**6**
	PTHR12411 (*CP1*)	2	1	3		1		2						
	PTHR22904 (*TPR* repeat)			3		1								
miR398	**PTHR10003 (*SOD*)**	**3**	**1**	**6**	**4**	**1**	**2**	**1**	**2**	**1**		**3**		**3**
	PTHR10122 (*COX*)	1		2		2			1					
	PTHR11709 (*LACCASE*)					1	2	1						
	PTHR33021 (*PLANTACYANIN*)	1				2			1					
	PTHR12483 (*COPPER TRANSPORTER; COPT*)			2	3		1							
	PTHR10593 (*SERINE/THREONINE-PROTEIN KINASE RIO; STPKR*)			6					2					
	PTHR11695 (*ALCOHOL DEHYDROGENASE; ADH*)			5		1	1							

Table 3.1. Distribution of HE target family members across species. The number of HE targets from the primary target family (in bold) and secondary target families found in each species. Dicots are highlighted in purple, monocots in blue and *A. trichopoda* in orange.

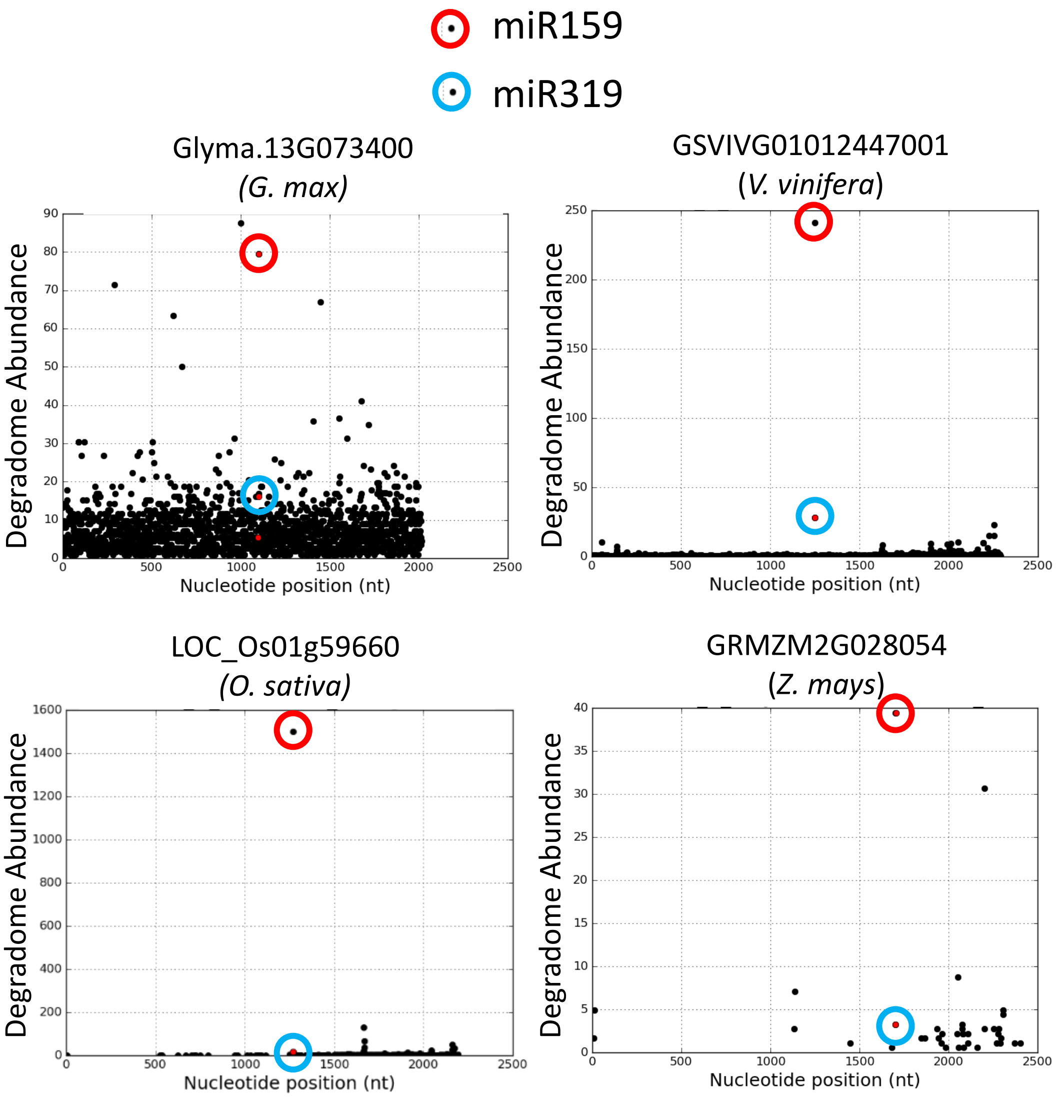


Figure 3.3. Cross regulation of miR319 and miR159 *MYB* gene targets. T-plots of the same *MYB* genes that were found to be HE targets for both miR159 and miR319. T-plots were taken from the same degradome libraries for both miRNAs in each species. Dots circled in red indicate the cleavage tag corresponding to cleavage by miR159, and the blue circle for miR319.

3.2.3 Target families of the same miRNA are commonly functionally related

Supporting the hypothesis that a secondary target would need to be compatible with the primary miRNA-target relationship (Li et al., 2014), three of six proposed secondary miRNA families were from functionally related processes to the primary target family. For instance, functional studies in Arabidopsis have shown that for the miR159 target families, *MYB* and *NOZZLE*, both are involved in anther development (Schiefthaler et al., 1999; Millar & Gubler, 2005); the miR395 targets, *ATP-SULFURYLASE* (*APS*) and *SULFATE TRANSPORTER2;1* (*SULTR2;1*), are both involved in sulphur metabolism and transport (Liang et al., 2010); and the miR398 targets, *SUPEROXIDE DISMUTASE* (*SOD*) and *CYTOCHROME C OXIDASE* (*COX*), are both involved in response to oxidative stress (Sunkar et al., 2006; Yamasaki et al., 2007) (Table 3.1). Furthermore, for miR398, *LACCASE* (*LAC*) and *PLANTACYANIN*, which are copper proteins like *SOD* were also identified as secondary target families (Abdel-Ghany & Pilon, 2008). An additional copper transporter gene family (*COPT*; PTHR12483) was also identified as a secondary target family of miR398 in dicots outside of Arabidopsis (Table 3.1) (Naya et al., 2014). Together, this suggests that in the instances in which a secondary target family is acquired, they are likely from functionally related processes.

3.2.4 Complementarity is not an absolute determinant of HE targets across miRNAs

Previous miRNA target prediction programs have relied heavily on the ranking of targets by miRNA-target complementarity (Bonnet et al., 2010; Dai et al., 2011; Sun et al., 2011; Dai et al., 2018). However, it is unclear how strict the correlation is, as there are targets with 3-5 mismatches that are strongly miRNA-regulated, whilst, there are genes with 0-2 mismatches which are poorly regulated (Brousse et al., 2014; Liu et al., 2014; Zheng et al., 2017).

Therefore, to investigate the extent that miRNA-target complementarity can be used as an indicator of HE targets, the average Expectation Score of all HE and LE targets was analysed (Figure 3.4). For most miRNAs, the average Expectation Scores were generally lower for the HE targets compared to the LE targets. For some miRNA targets, [eg. miR160, miR164, miR166, miR171, miR394, miR403 and tasiARF], HE targets had a much lower average Expectation Score compared to LE targets (≤ ½) suggesting mismatches are not tolerated. In these cases, low Expectation Scores may be a likely indicator of HE targets. However, the average Expectation Scores were not statistically different for all miRNAs [eg. miR162, miR319, miR395, and miR408] which suggests that it is not a reliable indicator for all miRNAs. Furthermore, Expectation Scores varied greatly within the HE and LE targets suggesting there are many exceptions where the Expectation Score of a target is not indicative of a HE. Further analyses of miRNAs using only their primary and secondary target families found no statistical difference in the average Expectation Score between HE and LE targets for most miRNAs (Figure 3.5). The average Expectation Scores were only found to be significantly lower in HE targets than LE targets for four miRNA-target families [ie. miR164:*NAC*, miR171:*HAM*: miR398:*SOD* and tasiARFs:*ARF*] suggesting that in these cases Expectation Scores may be a likely indicator of HE targets. For miR395:*APS*, the average Expectation Score of the HE targets was even significantly higher than the LE targets. Altogether, these results suggest that miRNA-binding site complementarity requirements vary greatly between each miRNA-target pair and in most cases the Expectation Score is not a reliable indicator of a HE target. As such, ranking the confidence of a gene as a miRNA target based on Expectation Score cannot be generally applied across miRNAs-target pairs. This implies factors additional to miRNA-binding site complementarity are involved in the miRNA-mediated regulation of a target.

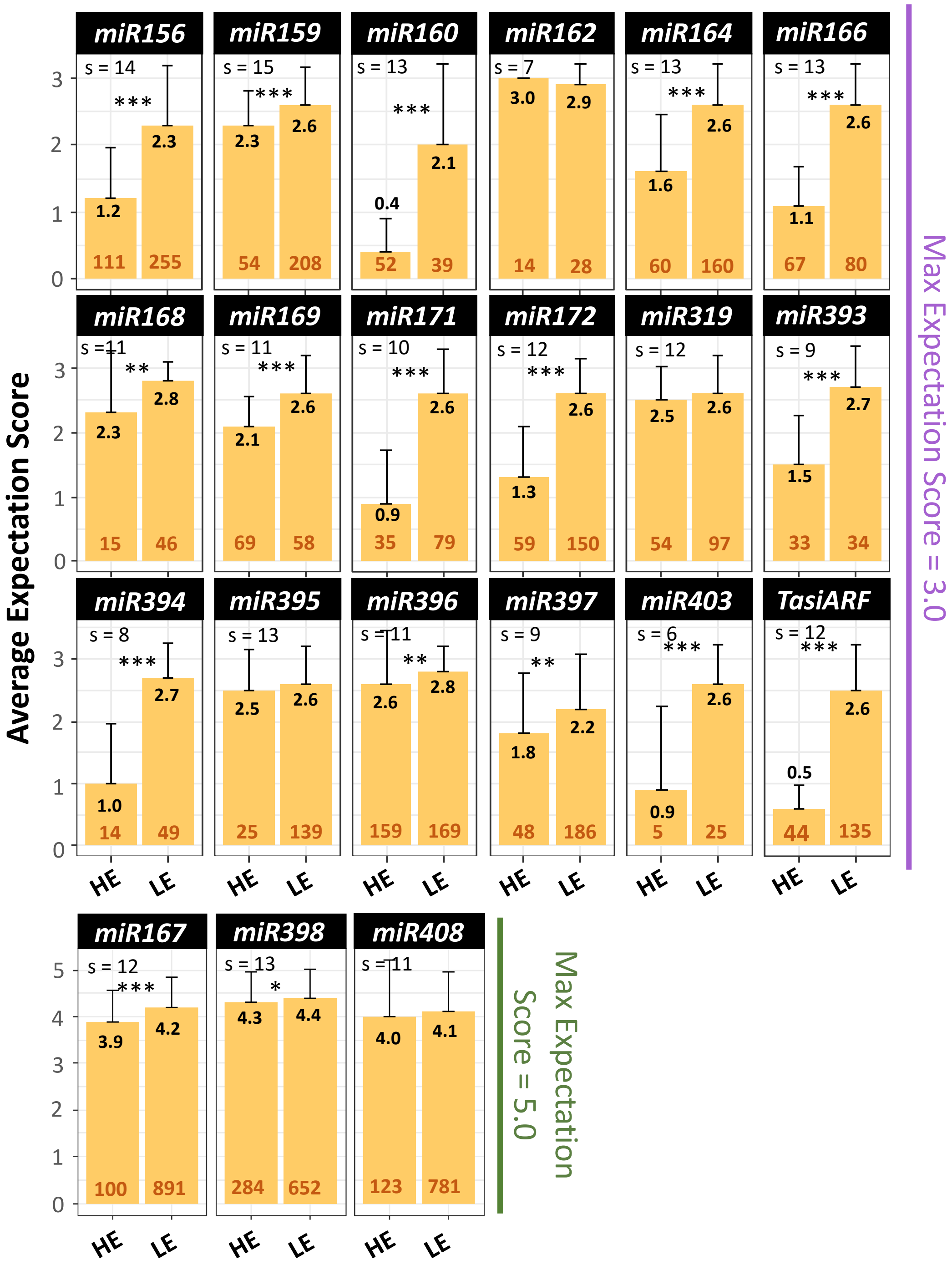


Figure 3.4. The average expectation score for HE and LE targets across species for each miRNA. The expectation score for target prediction with psRNATarget was increased to 5.0 (green) for miR167, miR398 and miR408 as the known targets for Arabidopsis exceed an expectation score of 3.0. Expectation Score for targets for all other miRNA do not exceed 3.0 (purple). Bolded numbers on and above bars indicate the average expectation score. ‘s’ indicates the number of species used for each analysis. Bolded orange numbers at the bottom of the bars indicate the number of genes analysed. Asterisks indicate statistical difference between the Expectation Scores of HE and LE targets where ‘*’ indicates P ≤ 0.05, ‘**’ indicates P ≤ 0.01, ‘***’ indicates P ≤ 0.001 and no asterisks indicates no significant difference.

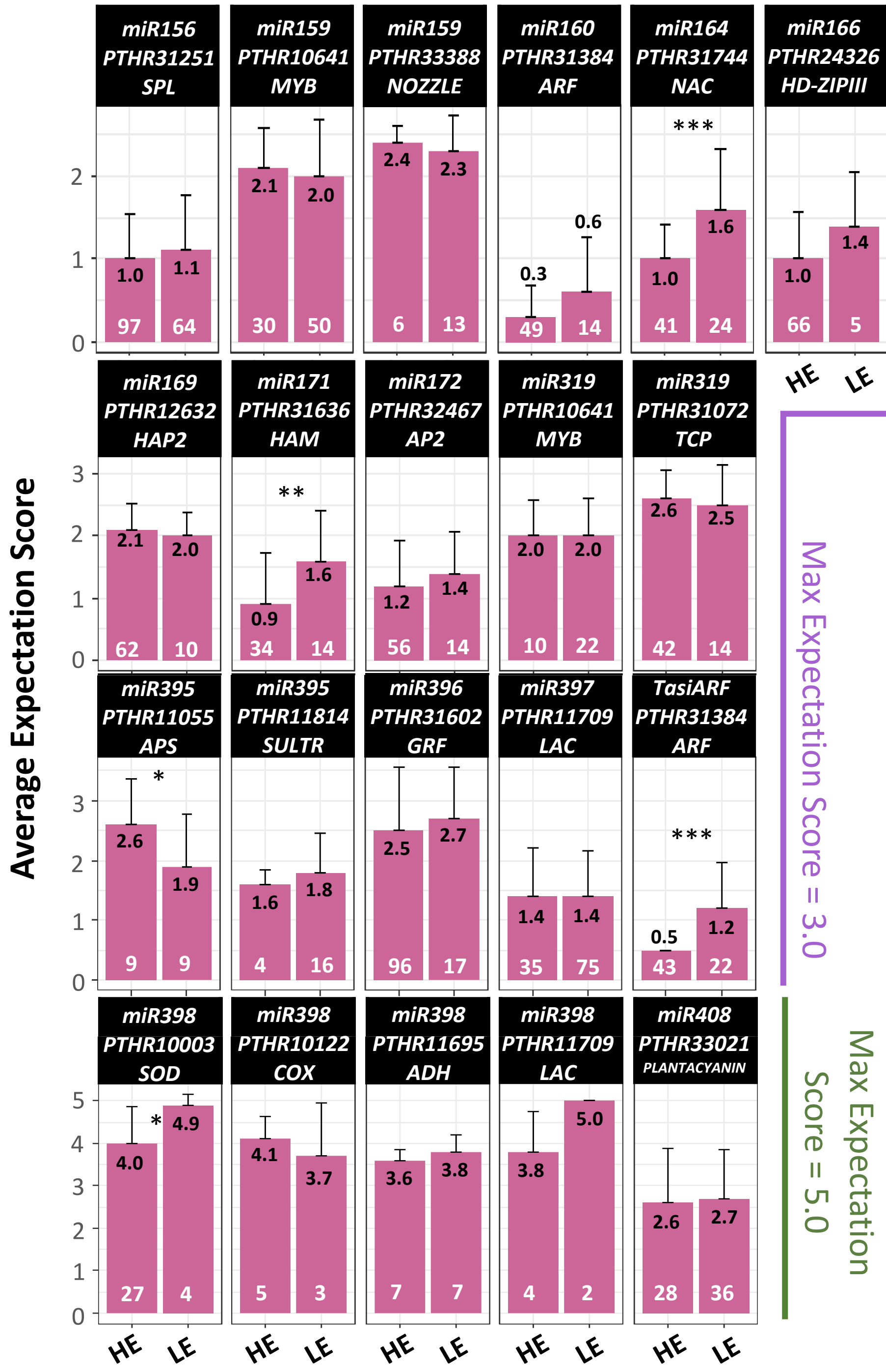


Figure 3.5. The average expectation score for HE and LE targets across species for miRNA:primary and secondary target family modules. Only targets from one primary or secondary target family were analysed in each box. The expectation score for target prediction with psRNATarget was increased to 5.0 (green) for miR167, miR398 and miR408 as the known targets for Arabidopsis exceed an expectation score of 3.0. Expectation Score for targets for all other miRNA do not exceed 3.0 (purple). Bolded black numbers on and above bars indicate the average expectation score. Bolded white numbers at the bottom of the bars indicate the number of genes analysed. Asterisks indicate statistical difference between the Expectation Scores of HE and LE targets where ‘*’ indicates P ≤ 0.05, ‘**’ indicates P ≤ 0.01, ‘***’ indicates P ≤ 0.001 and no asterisks indicates no significant difference.

3.2.5 Conserved nucleotides flanking the miR159-binding site in *MYB* homologues correlate with HE targets across species

As previously mentioned, highly conserved RNA secondary structures that are associated with the miR159-binding site in *MYB33* promote miR159-mediated regulation in Arabidopsis (Zheng et al., 2017). Sequence alignments had shown conserved flanking nucleotide sequences that corresponded to the stems of these secondary structures were present in *MYB33* and *MYB65*, the two *MYB* homologues that were strongly regulated by miR159 (Zheng et al., 2017). Therefore, are these conserved flanking sequencings a feature characteristic of strong miR159-mediated regulation across diverse species?

Firstly, to quantitatively define these flanking features, the conservation of the sequences was measured from an alignment of *MYB33* homologues using the program, phyloP (phylogenetic *P*-values; Pollard et al., 2010). Results found five sequences with four or more consecutive nucleotides undergoing slower nucleotide substitution rates than expected under neutral drift compared to neighbouring nucleotides, even at wobble positions (FDR-adjusted phyloP score ≥ 1.0) (Figure 3.6 A-C). The RNA secondary structure was predicted for this aligned consensus sequence using RNAalifold (Bernhart et al., 2008), which generated an RNA secondary structure consistent with that previously reported (Figure 3.6 D; Zheng et al., 2017). The RNA secondary structure consisted of two stem-loops, with consistent spacing and conformation across species that have been designated stem-loop 1 (SL1) and 2 (SL2) (Zheng et al., 2017) (Figure 3.6 A). These conserved sequences and their spacings to one another were used as the criteria to identify *MYB* homologues with this predicted RNA secondary structure feature (Figure 3.7 A).

Strikingly, analysis of HE and LE *MYB* targets found these conserved sequences only occurred in HE targets, with 23 of the 30 *MYB* HE homologues possessing these sequence features (Figure 3.7 B). Clearly, these conserved sequences are highly correlated with HE *MYB* targets across diverse species, including the ancient basal angiosperm, *Amborella trichopoda*. This strongly supports the idea that an ancient RNA secondary structural element has been central in the miR159-mediated regulation of *MYB* targets across species (Table 3.3).

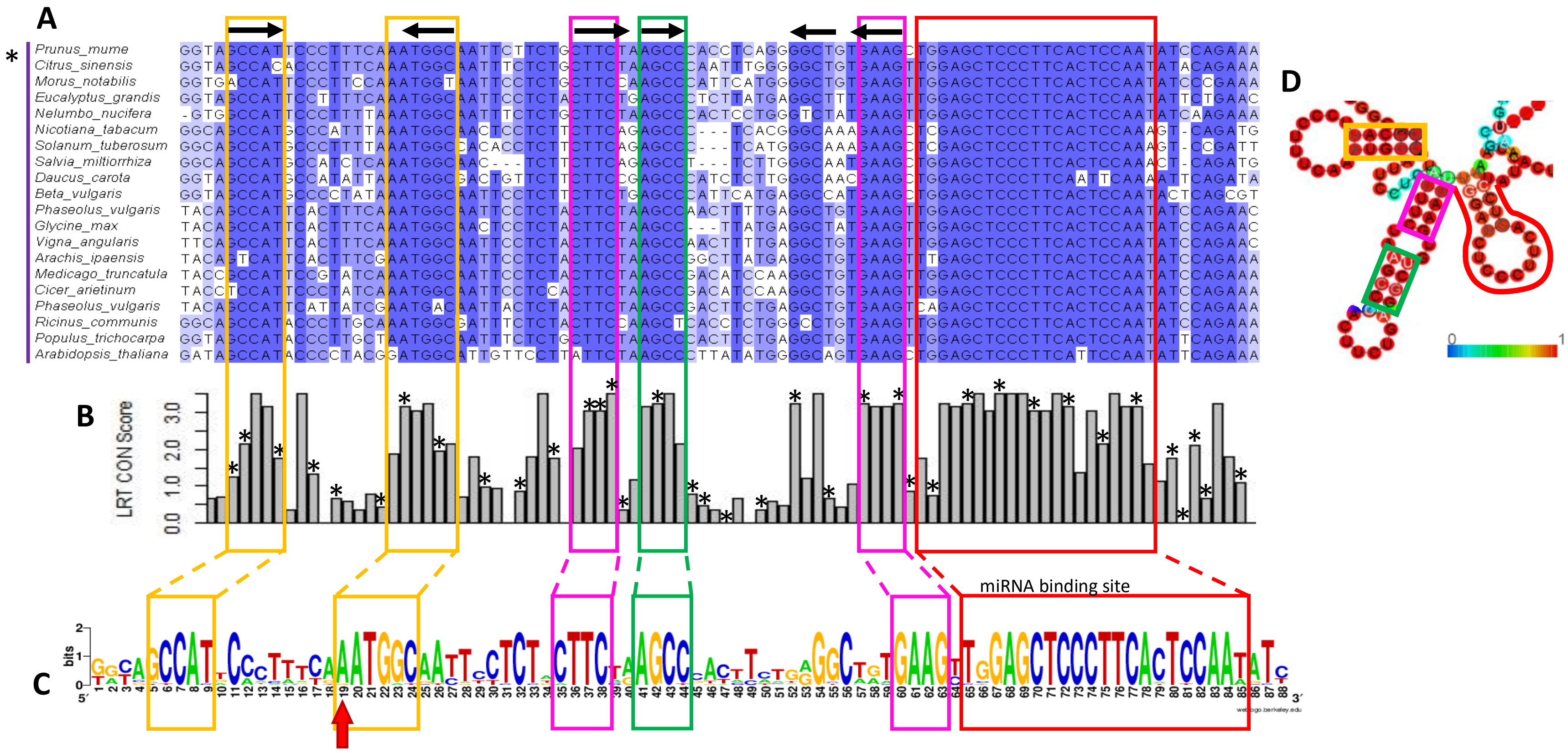


Figure 3.6. Conserved sequences flanking the binding site of miR159:*MYB33* homologues. A) MSA constructed from twenty *MYB33* homologues from various dicots. The miR159 binding site is indicated by a red box, and the conserved flanking sequences in green, pink, blue and/or yellow boxes in A to D. Black arrows above the MSA and matching coloured boxes around conserved flanking sequences indicate base pairing from the predicted secondary structure in D. Asterisks indicate plant classification where '*' indicates dicots. B) The phyloP score at each nucleotide position around the miR159:MYB binding site. '*' denote a degenerate nucleotide site in reference to the *Arabidopsis thaliana* sequence. C) The sequence logo of the binding site and conserved flanking sequences. The height of each nucleotide indicates its relative frequency at this postion. The red arrow indicates an additional conserved nucleotide to those found in Zheng et al. (2017). D) The secondary structure and the probability of base pairing predicted from the consensus sequence from the MSA.

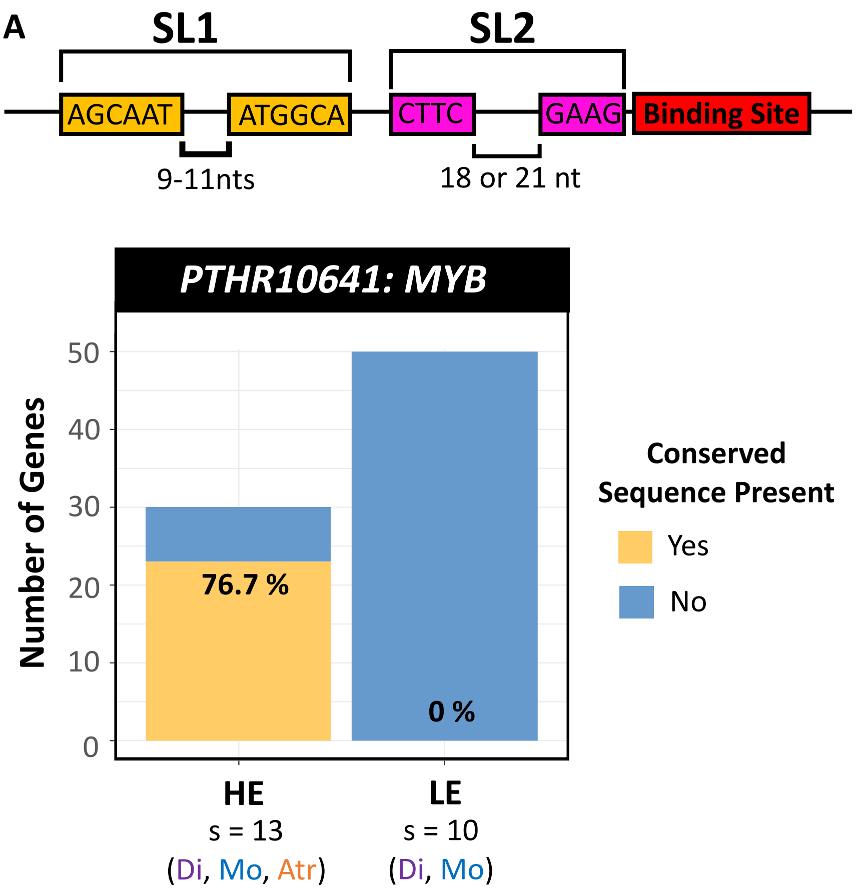


Figure 3.7. *MYB* family members with conserved sequences flanking the miR159 binding site. A) A schematic of the conserved sequences flanking the binding site and their locations as determined by MSA for HE and LE *MYB* targets. Conserved sequences were considered present if HE targets possessed sequences predicted to form either the first or second stem-loop (SL1 and SL2, respectively). Additionally, the conserved sequences for SL1 must be 9 to 11 nts apart to be considered present as consistently observed across gene homologues. SL2 sequences must be 18 or 21 nts apart (further detailed in Figure S3). B) HE and LE targets with conserved sequences present are indicated in yellow and targets without conserved sequences are indicated in blue. Numbers above bars indicate the total number of targets analysed. The number of HE and LE targets with the conserved sequence out of total genes is expressed as a percentage in the yellow bars. '*spp*' indicates the number of species the targets are from.

3.2.6 Multiple conserved target families have conserved sequences flanking their miRNA binding sites

As conserved sequences flanking the miR159-binding site were found to be a feature characteristic of strongly regulated *MYB* homologues across diverse species, this raised the possibility that similar scenarios may have arisen in other highly conserved target families over evolutionary time. To investigate this, multiple sequence alignments (MSAs) were performed on the primary target families, followed by phyloP analysis to identify whether nucleotide conservation extends beyond their miRNA-binding site. Ten targets were identified with conserved sequences extending beyond the binding site (Figure 3.8 - 3.21). Conservation was present even at wobble positions and, in some cases, comparable to the highly conserved binding site. A sequence was considered conserved as defined in the Materials and Methods.

miR160:*ARF10* – *ARF10* homologues were identified in multiple dicot and monocot species and twenty were used to construct a MSA (Figure 3.8). The miR160 binding site was invariant in all but four homologues. Conservation extended six nucleotides beyond the miR160 binding site at both the 5` and 3` ends (Figure 3.8). Both these flanking sequences contained four nts that are complementary to the miR160-binding site, hence potentially forming a conserved RNA secondary structure incorporating the miR160-binding site (Figure 3.8 D).

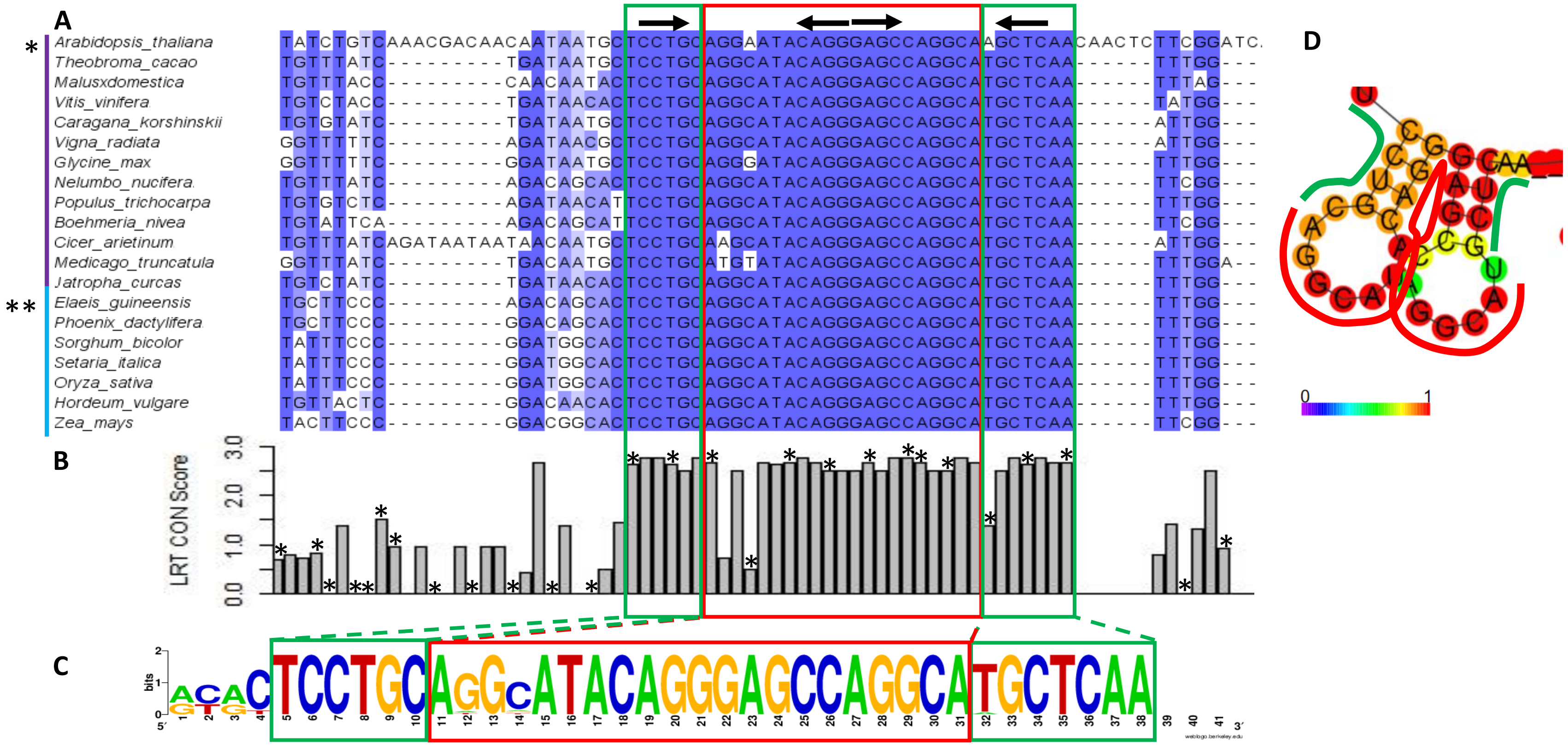


Figure 3.8. Conserved nucleotides flanking the binding site of miR160:*ARF10* homologues. A) MSA constructed from twenty *ARF10* homologues from various dicots and monocots. The miR160 binding site is indicated by a red box, and the conserved flanking sequences in a green box in A to D. Black arrows above the MSA indicate base pairing from the predicted secondary structure in D. Asterisks indicate plant classification where '*' indicates dicots, and '**' monocots. B) The phyloP score at each nucleotide position around the binding site. '*' denote a degenerate nucleotide site in reference to the *Arabidopsis thaliana* sequence. C) The sequence logo of the binding site and conserved flanking sequences. The height of each nucleotide indicates its relative frequency at this position. D) The secondary structure and the probability of base pairing predicted from the consensus sequence from the MSA.

miR160:*ARF17* – *ARF17* homologues were identified from dicots, monocots, the basal angiosperm, *A. trichopoda*, lycophytes, and to the oldest extant lineage of land plants, bryophytes. Aside from the 5`-nucleotide position, the miR160 binding site was invariant in all but three homologues and twenty were used to construct a MSA (Figure 3.9). Conservation extended to seven flanking nucleotides directly upstream of the miR160 binding site (Figure 3.9) and was near-identical in homologues across these diverse lineages, suggesting it corresponds to an ancient motif. Interestingly, all seven conserved flanking nucleotides are complementary to the binding site and are predicted to form a conserved secondary structure at a high base-pair probability (Figure 3.9 D).

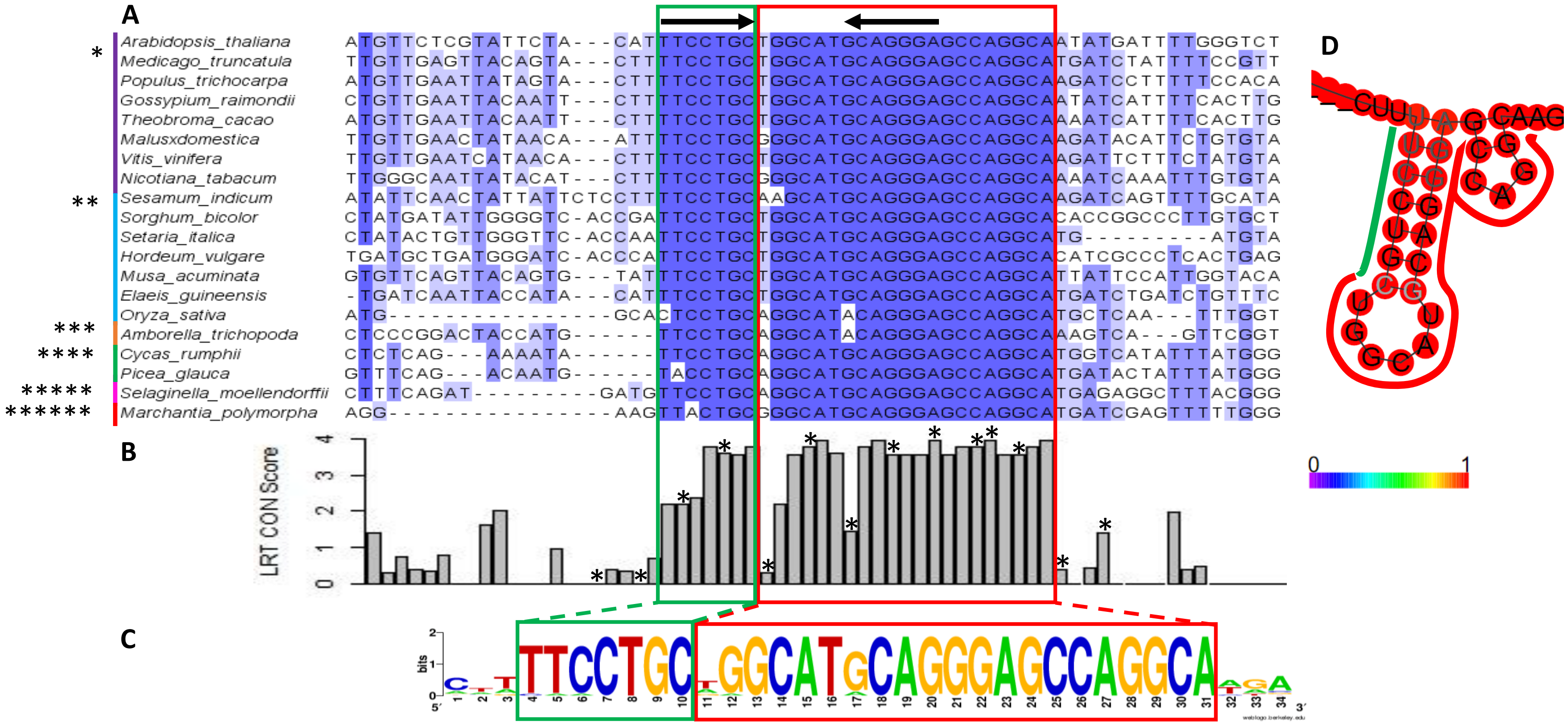
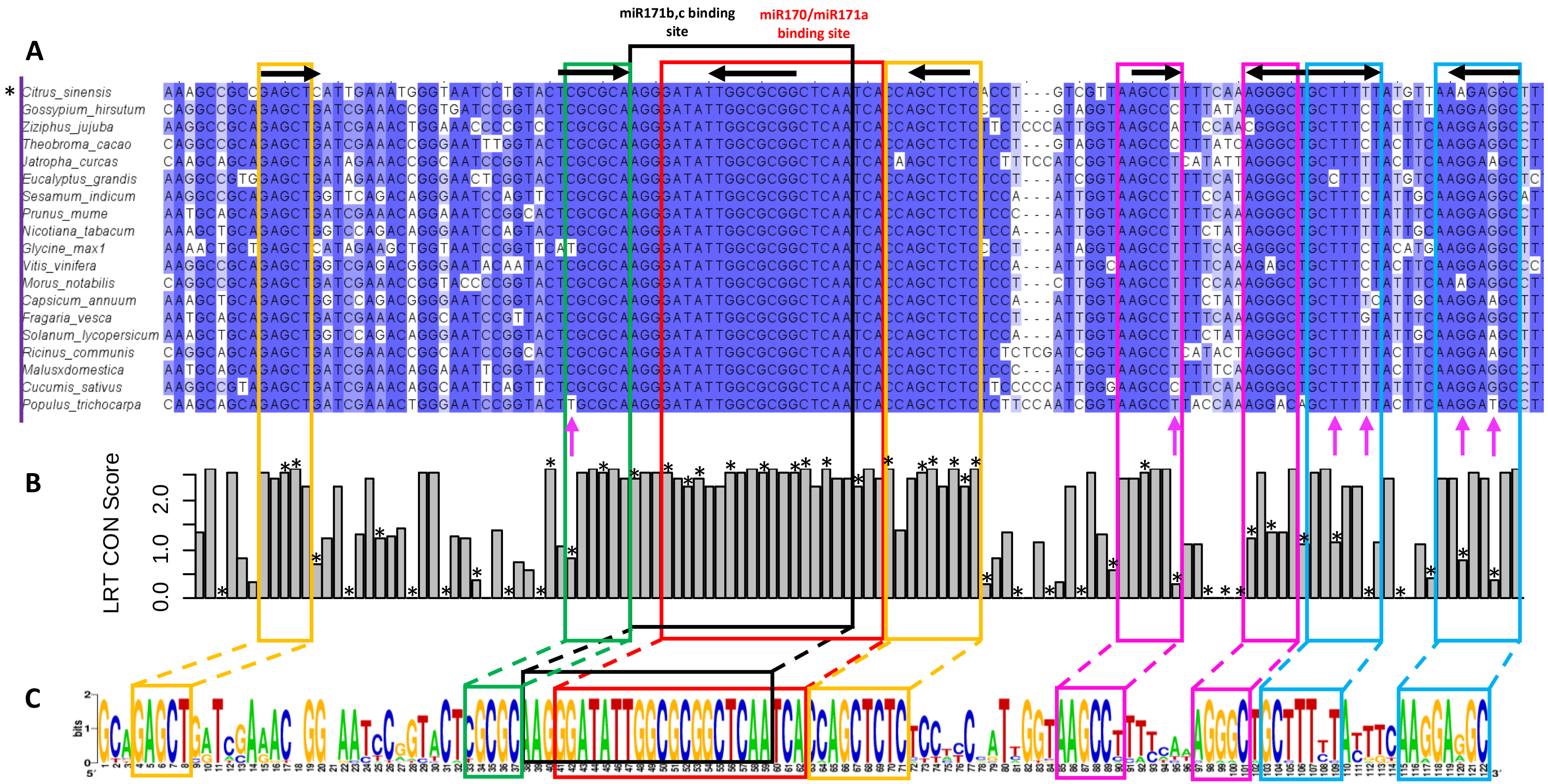


Figure 3.9. Conserved nucleotides flanking the binding site of miR160:*ARF17* homologues. A) MSA constructed from twenty *ARF17* homologues from various lineages of land plants. The miR160 binding site is indicated by a red box, and the conserved flanking sequences in a green box in A to D. Black arrows above the MSA indicate base pairing from the predicted secondary structure in D. Asterisks indicate plant classification where '*' indicates dicots, '**' monocots, '***' *A. trichopoda*, '****' gymnosperms, '*****' lycophytes, and '******' bryophytes. B) The phyloP score at each nucleotide position around the binding site. '*' denote a degenerate nucleotide site in reference to the *Arabidopsis thaliana* sequence. C) The sequence logo of the binding site and conserved flanking sequences. The height of each nucleotide indicates its relative frequency at this position. D) The secondary structure and the probability of base pairing predicted from the consensus sequence from the MSA.

miR171:*HAM1* – *HAM1* homologues [also known as *SCARECROW-like*] were identified from multiple dicot species and twenty were used to construct a MSA (Figure 3.10). *HAM1* has two overlapping binding sites (miR171b,c, and miR170/miR171a) (Bari et al., 2013). No nucleotide variation was found in either binding sites across all homologues analysed. Analysing the sequences directly upstream and downstream of the miRNA binding sites found seven conserved sequences ranging from five to nine nucleotides long (Figure 3.12 A – C). All conserved sequences had nucleotides complementary to another of these conserved sequences or to the binding site and were predicted to form the stems of RNA secondary structures with varying probability (Figure 3.10E). The most distal conserved sequences downstream of the binding site were predicted to form two conserved stem loops. Although, these sequences showed nucleotide variations, these were still compatible with base pairing and therefore were still predicted to form the two stem loops (Figure 3.10 E). Interestingly, like *ARF10* and *ARF17*, the conserved sequence directly upstream of the binding sites was predicted to base pair with a GC rich sequence in the binding site to form a strong stem (Figure 3.10 D).


A
miR171b,c binding site
miR170/miR171a binding site
*
Citrus_sinensis
Gossypium_hirsutum
Ziziphus_jujuba
Theobroma_cacao
Jatropha_curcas
Eucalyptus_grandis
Sesamum_indicum
Prunus_mume
Nicotiana_tabacum
Glycine_max1
Vitis_vinifera
Morus_notabilis
Capsicum_annuum
Fragaria_vesca
Solanum_lycopersicum
Ricinus_communis
Malusxdomestica
Cucumis_sativus
Populus_trichocarpa
B
LRT CON Score
0.0
1.0
2.0
C
bits
2
1
0
5'
3'

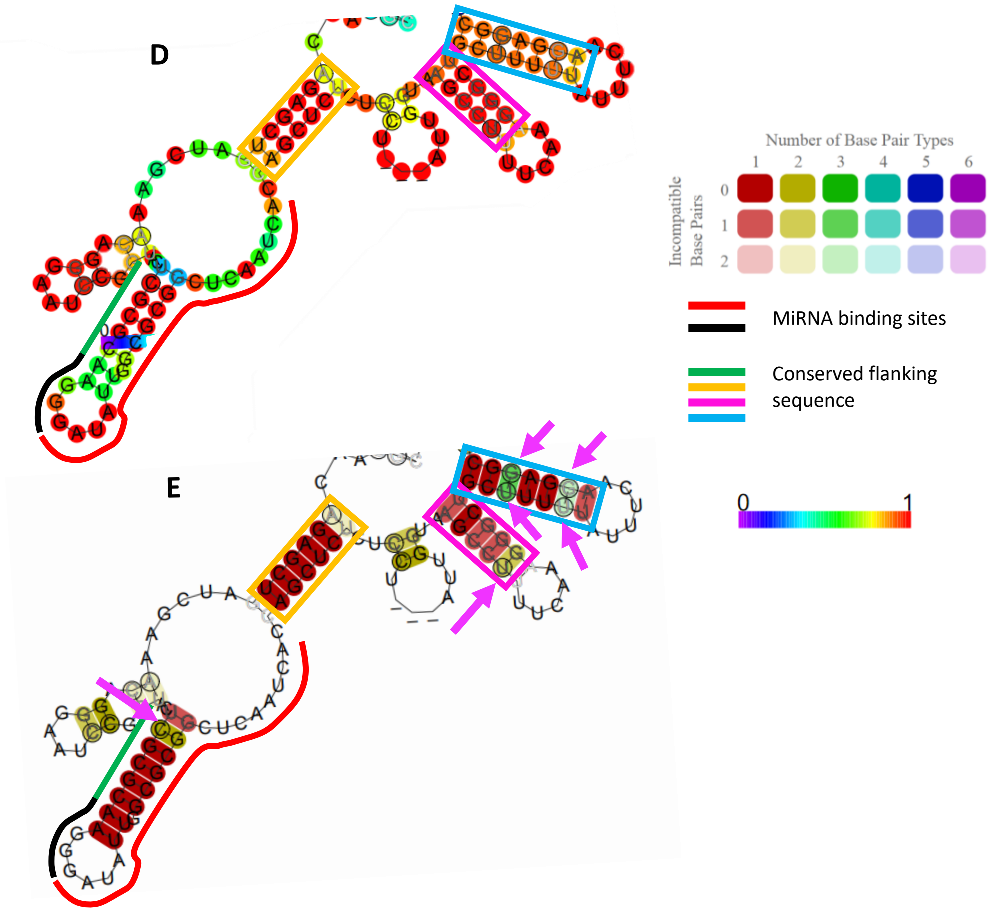


Figure 3.10. Conserved nucleotides flanking the binding site of miR171:*HAM1* homologues. A) MSA constructed from twenty *HAM* homologues from various dicots. *HAM1* homologues have two binding sites which overlap by three nucleotides. The miR171b,c binding site is indicated by a black box, and the miR170/miR171a in red. The conserved flanking sequences are indicated by green, pink, blue and/or yellow boxes in A to E. Black arrows above the MSA indicate base pairing from the predicted secondary structure in D and E. Asterisks indicate plant classification where '*' indicates dicots. Pink arrows indicate positions with nucleotide variations which are still compatible with base pairing and were considered conserved despite a lower phyloP score. B) The phyloP score at each nucleotide position around the binding site. '*' denote a degenerate nucleotide site in reference to the *Arabidopsis thaliana* sequence. C) The sequence logo of the binding site and conserved flanking sequences. The height of each nucleotide indicates its relative frequency at this position. D) The secondary structure and the probability of base pairing predicted from the consensus sequence from the MSA. E) The predicted secondary structure showing conservation annotation. Colours represent the number of base pairs types (ie. AU, UA, CG, GC, UG, GU), and hue the number of non-conserved nucleotides at that position.

miR319:*TCP2* –Twenty *TCP2* homologues were identified in multiple dicot and monocot species and used to construct an MSA (Figure 3.11). Nine nucleotides were found to be conserved upstream of the miR319 binding site, five of which corresponded to wobble positions, suggesting conservation at the RNA level. This conserved sequence was not predicted to form an RNA secondary structure.

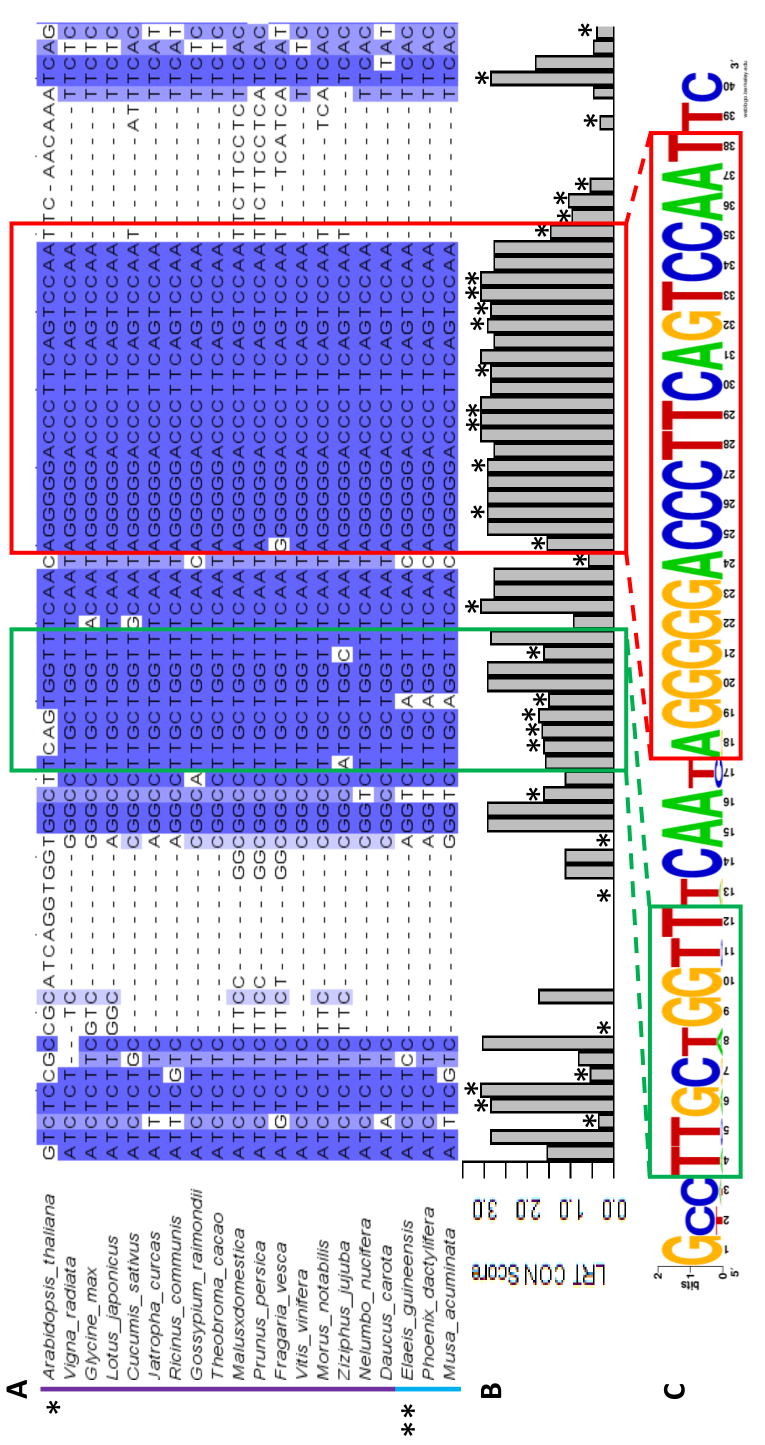


Figure 3.11. Conserved nucleotides flanking the binding site of miR319:*TCP2* homologues. A) MSA constructed from twenty *TCP2* homologues from various dicots and monocots. The miR319 binding site is indicated by a red box, and the conserved flanking sequences in a green box in A to D. Black arrows above the MSA indicate base pairing from the predicted secondary structure in D. Asterisks indicate plant classification where '*' indicates dicots, and '**' monocots. B) The phyloP score at each nucleotide position around the binding site. '*' denote a degenerate nucleotide site in reference to the *Arabidopsis thaliana* sequence. C) The sequence logo of the binding site and conserved flanking sequences. The height of each nucleotide indicates its relative frequency at this position.

miR319:*TCP4* –Similarly, for *TCP4*, twenty homologues from multiple dicot and monocot species were used to construct an MSA (Figure 3.12). In this case, two conserved sequences were found with one conserved sequence consisting of five nucleotides directly flanking the 5′ end of the binding site. Another sequence was downstream of the binding site and was four nucleotides long. Similar to *TCP2*, no RNA secondary structure was predicted to form from these sequences.

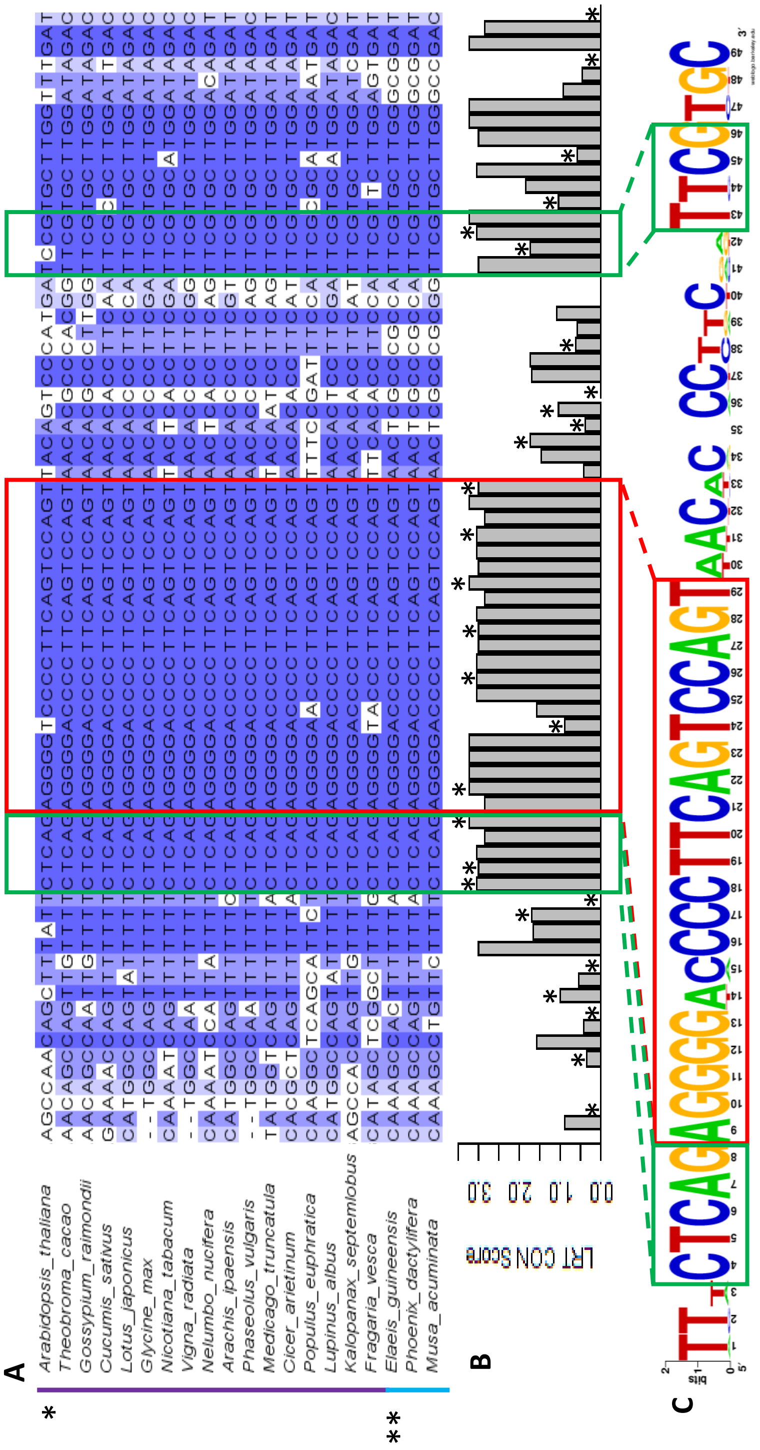


Figure 3.12. Conserved nucleotides flanking the binding site of miR319:*TCP4* homologues. A) MSA constructed from twenty *TCP4* homologues from various dicots and monocots. The miR319 binding site is indicated by a red box, and the conserved flanking sequences in a green box in A to D. Black arrows above the MSA indicate base pairing from the predicted secondary structure in D. Asterisks indicate plant classification where '*' indicates dicots, and '**' monocots. B) The phyloP score at each nucleotide position around the binding site. '*' denote a degenerate nucleotide site in reference to the *Arabidopsis thaliana* sequence. C) The sequence logo of the binding site and conserved flanking sequences. The height of each nucleotide indicates its relative frequency at this position.

miR390:*TAS3* – for the non-coding RNA *TRANS-ACTING SHORT INTERFERING RNA 3* (*TAS3*), a MSA of twenty *TAS3* homologues from dicot species found conservation to extend upstream of the miR390-binding site (Figure 3.13). These consisted of two conserved sequences which were five and nine nucleotides long (from distally to proximally of the binding site, respectively) and were separated by a three-nucleotide gap. These conserved sequences were not predicted to form an RNA secondary structure.

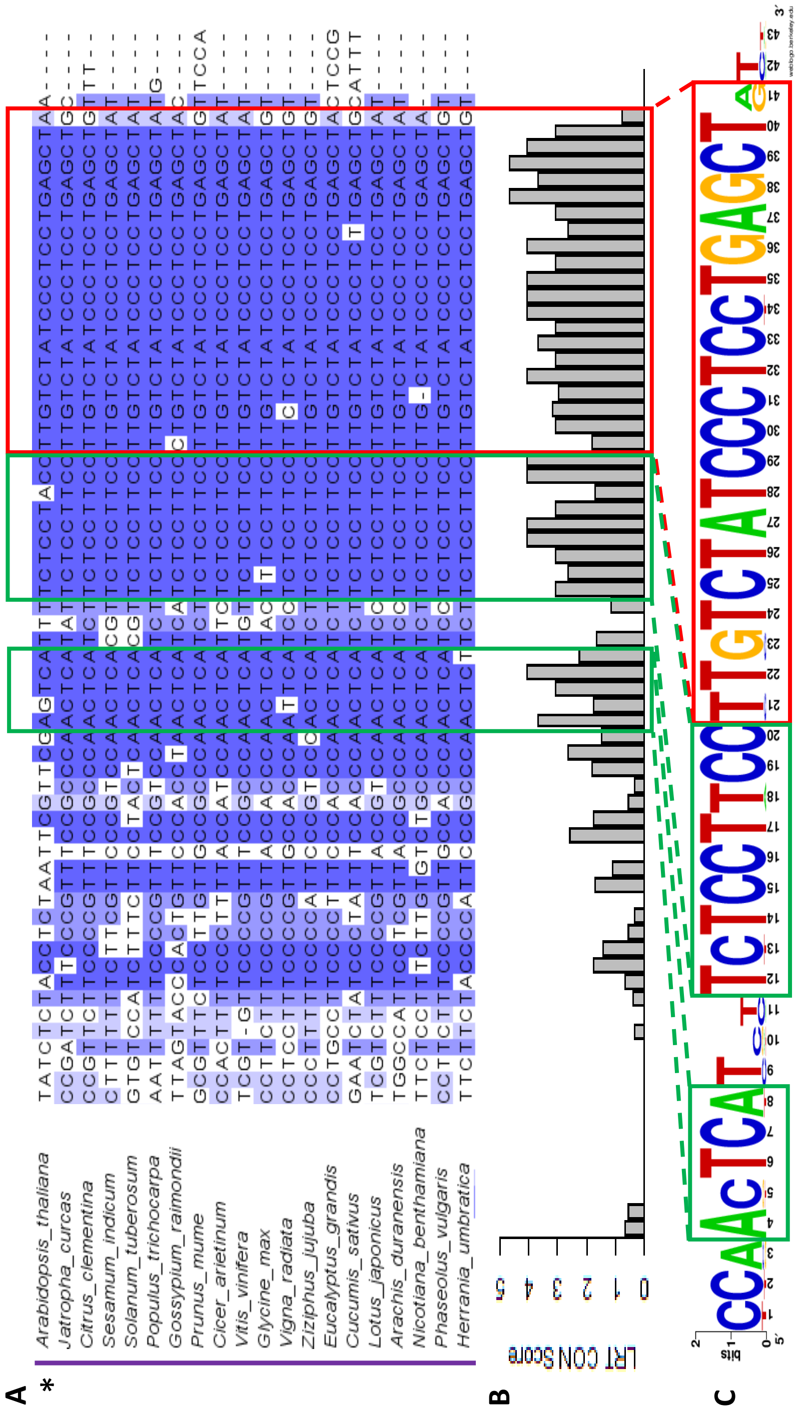


Figure 3.13. Conserved nucleotides flanking the binding site of miR390:TAS3 homologues. A) MSA constructed from twenty *TAS3* homologues from various dicots. The miR390 binding site is indicated by a red box, and the conserved flanking sequences in a green box in A to D. Black arrows above the MSA indicate base pairing from the predicted secondary structure in D. Asterisks indicate plant classification where '*' indicates dicots. B) The phyloP score at each nucleotide position around the binding site. '*' denote a degenerate nucleotide site in reference to the *Arabidopsis thaliana* sequence. C) The sequence logo of the binding site and conserved flanking sequences. The height of each nucleotide indicates its relative frequency at this position.

miR396:*GRF3* – Twenty *GROWTH-REGULATING FACTOR 3* (*GRF3*) homologues from multiple dicot and monocots species as well as *Amborella trichopoda* were used to construct a MSA (Figure 3.14). Three conserved sequences were identified upstream of the miR396 binding site but were not predicted to form an RNA secondary structure.

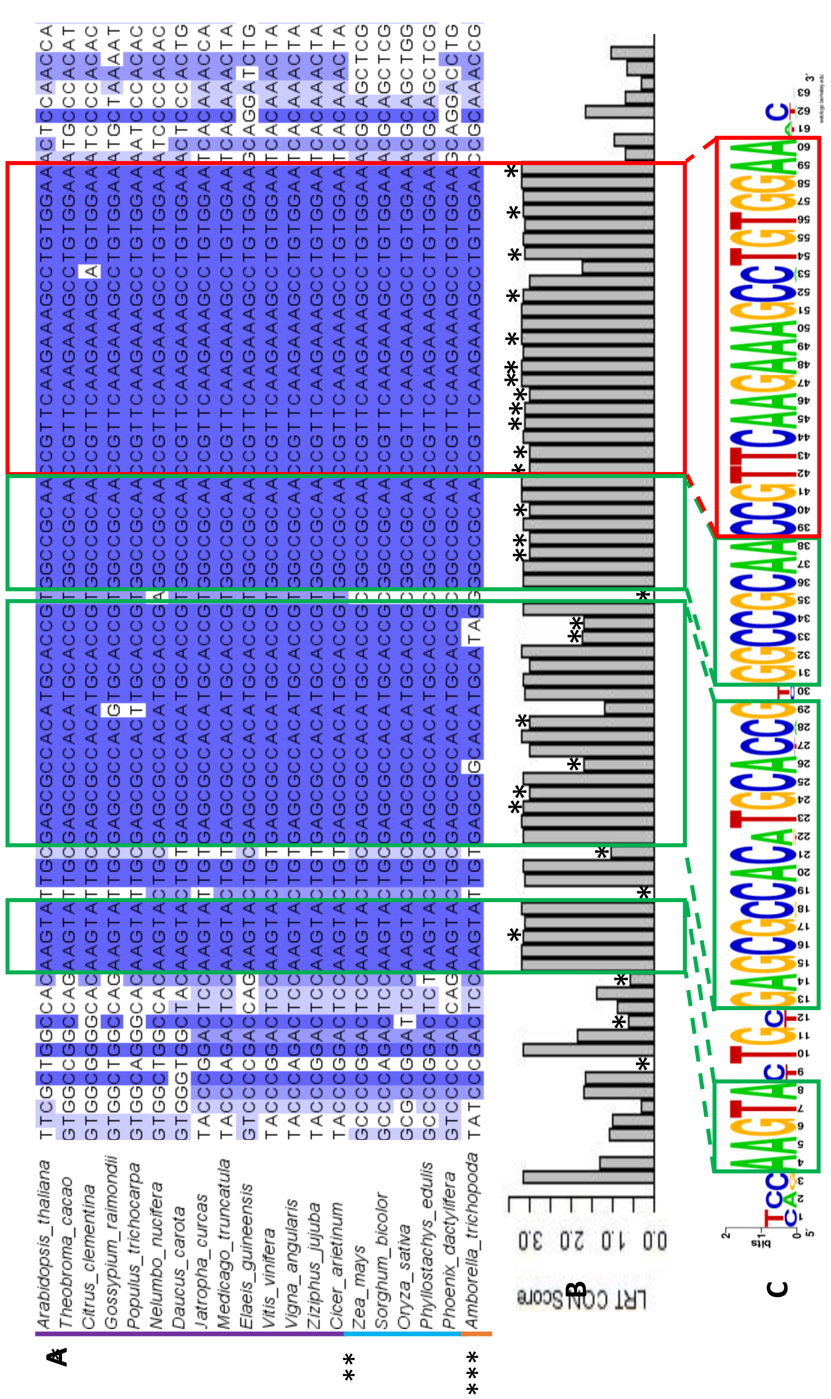


Figure 3.14. Conserved nucleotides flanking the binding site of miR396:GRF homologues. A) MSA constructed from twenty *GRF* homologues from various dicots. The miR396 binding site is indicated by a red box, and the conserved flanking sequences in a green box in A to D. Black arrows above the MSA indicate base pairing from the predicted secondary structure in D. Asterisks indicate plant classification where '*' indicates dicots, '**' monocots, '***' *A. trichopoda*. B) The phyloP score at each nucleotide position around the binding site. '*' denote a degenerate nucleotide site in reference to the *Arabidopsis thaliana* sequence. C) The sequence logo of the binding site and conserved flanking sequences. The height of each nucleotide indicates its relative frequency at this position.

miR399:*IPS1* – Like *TAS3*, *INDUCED BY PHOSPHATE STARVATION1* (*IPS1*) is also a non-coding RNA. From a MSA of eight *Brassicaceae* homologues, the miR399 binding site was found to be highly conserved except at the position coinciding with the central nucleotide in the 3-nt bulge required to inhibit miR399-guided cleavage (Franco-Zorrilla et al., 2007) (Figure 3.15). Conservation extended for 31 nucleotides downstream without any nucleotide variations. This conserved sequence was also predicted at high probability to form an RNA stem-loop (Figure 3.15 A – D). In a MSA of seven *IPS1* homologues from monocots, conservation was found to extend to 11 nucleotides upstream of the miR399 binding site but was not predicted to form an RNA secondary structure (Figure 3.15 E – G).

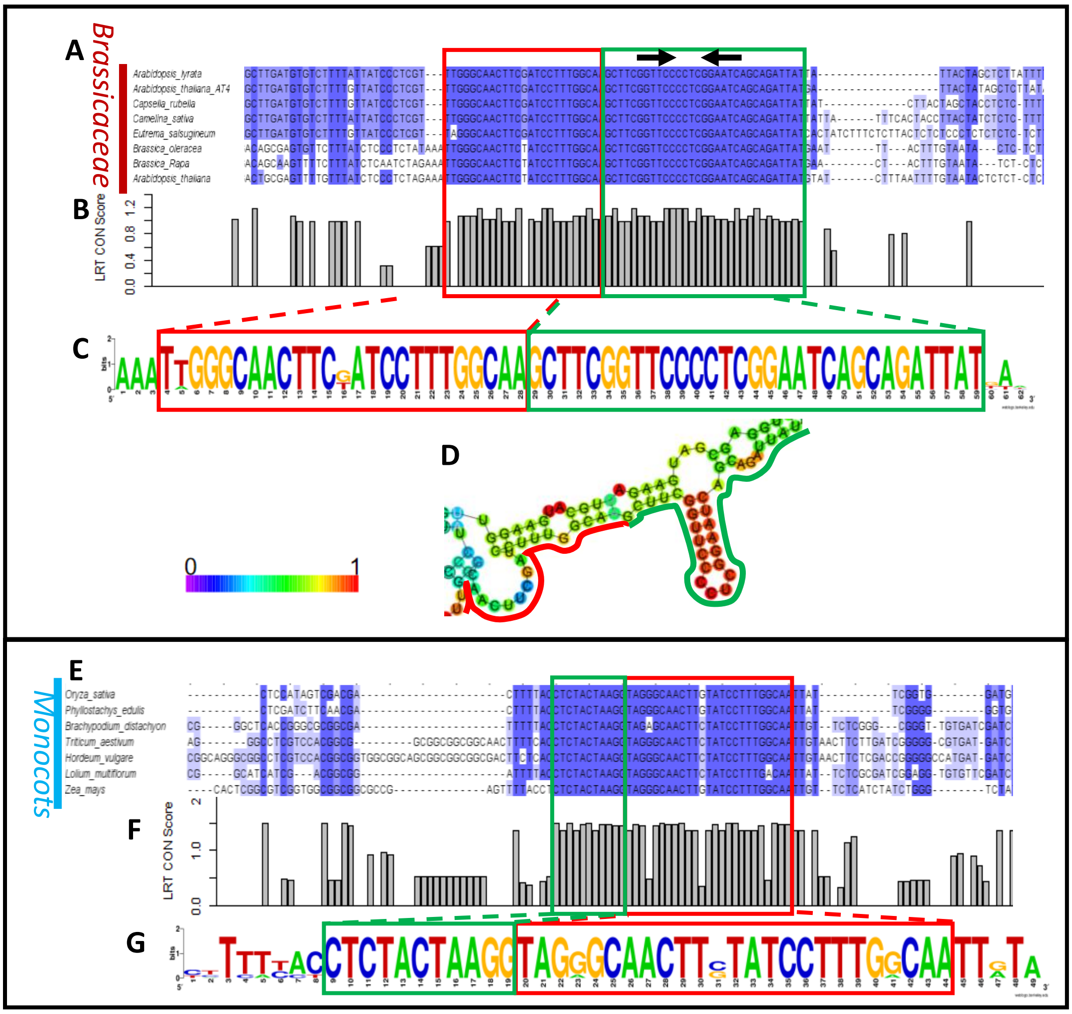
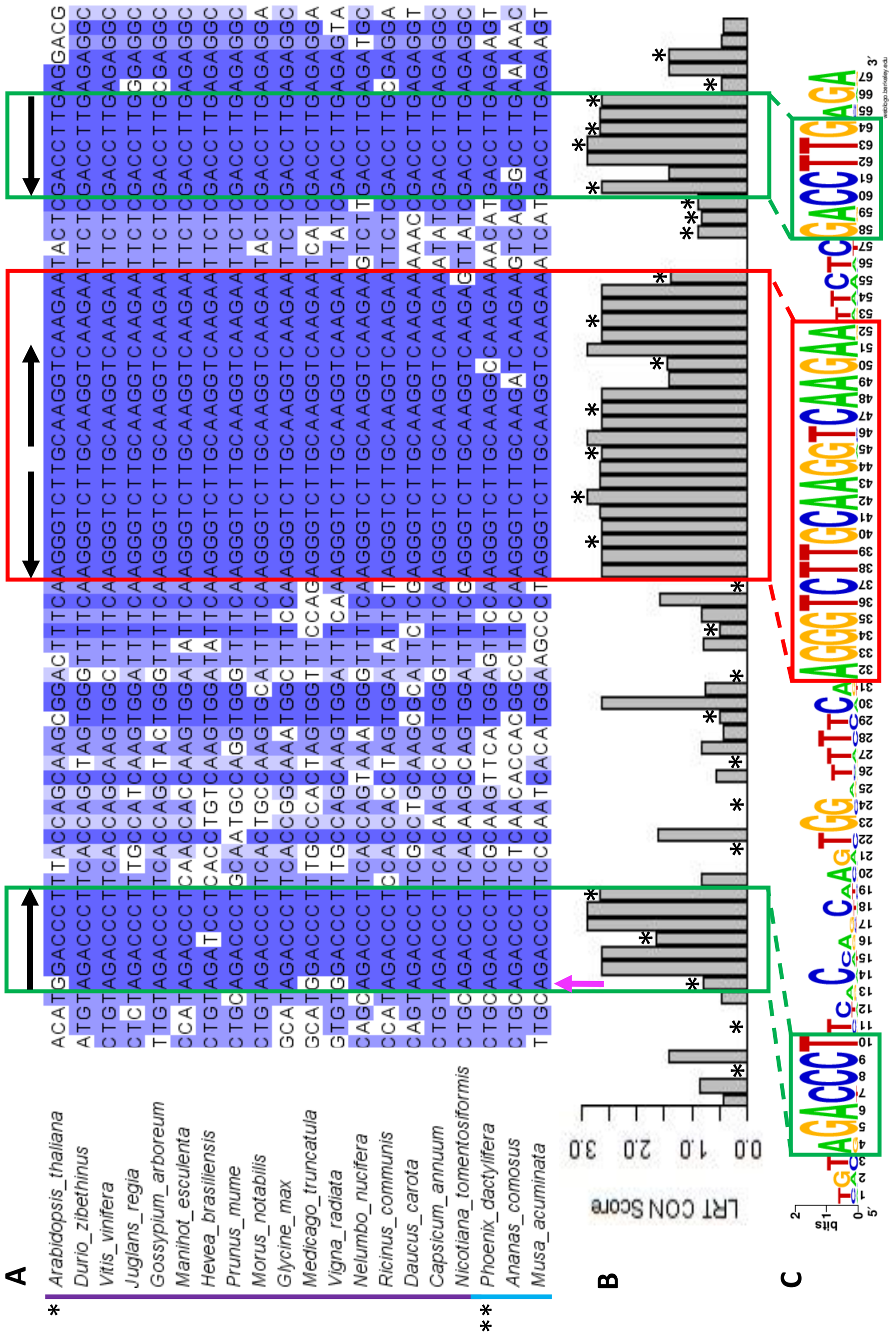


Figure 3.15. Conserved nucleotides flanking the binding site of miR399:*IPS1* homologues. A) MSA constructed from eight *IPS1* homologues from *Brassicaceae*. The miR399 binding site is indicated by a red box, and the conserved flanking sequences in a green box in A to D. Black arrows above the MSA indicate base pairing from the predicted secondary structure in D. B) The phyloP score at each nucleotide position around the binding site. C) The sequence logo of the binding site and conserved flanking sequences. The height of each nucleotide indicates its relative frequency at this position. D) The secondary structure and the probability of base pairing predicted from the consensus sequence from the MSA. E) MSA constructed from seven *IPS1* homologues from monocots. The miR399 binding site is indicated by a red box, and the conserved flanking sequences in a green box in E to G. Black arrows above the MSA indicate base pairing from the predicted secondary structure in D. F) The phyloP score at each nucleotide position around the binding site. G) The sequence logo of the binding site and conserved flanking sequences. The height of each nucleotide indicates its relative frequency at this position.

tasiARF:*ARF2* – Twenty *ARF2* homologues from multiple dicot and monocot species were used to construct a MSA (Figure 3.16). Two conserved sequences consisting of seven nucleotides each were identified upstream and downstream of the tasiARF binding site (Figure 3.16 A – C). All conserved nucleotides were complementary to the binding site and were predicted to base-pair at a high probability forming two stem-loops (Figure 3.16 A & D). Nucleotide variations were found in one position in the conserved sequences which was still compatible with base-pairing to the binding site.



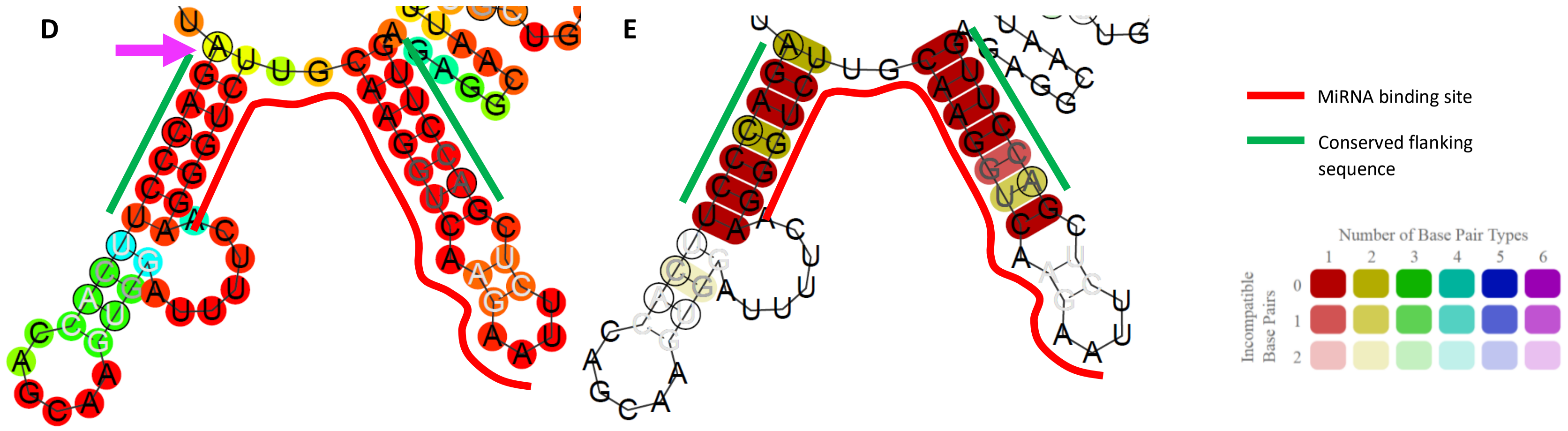
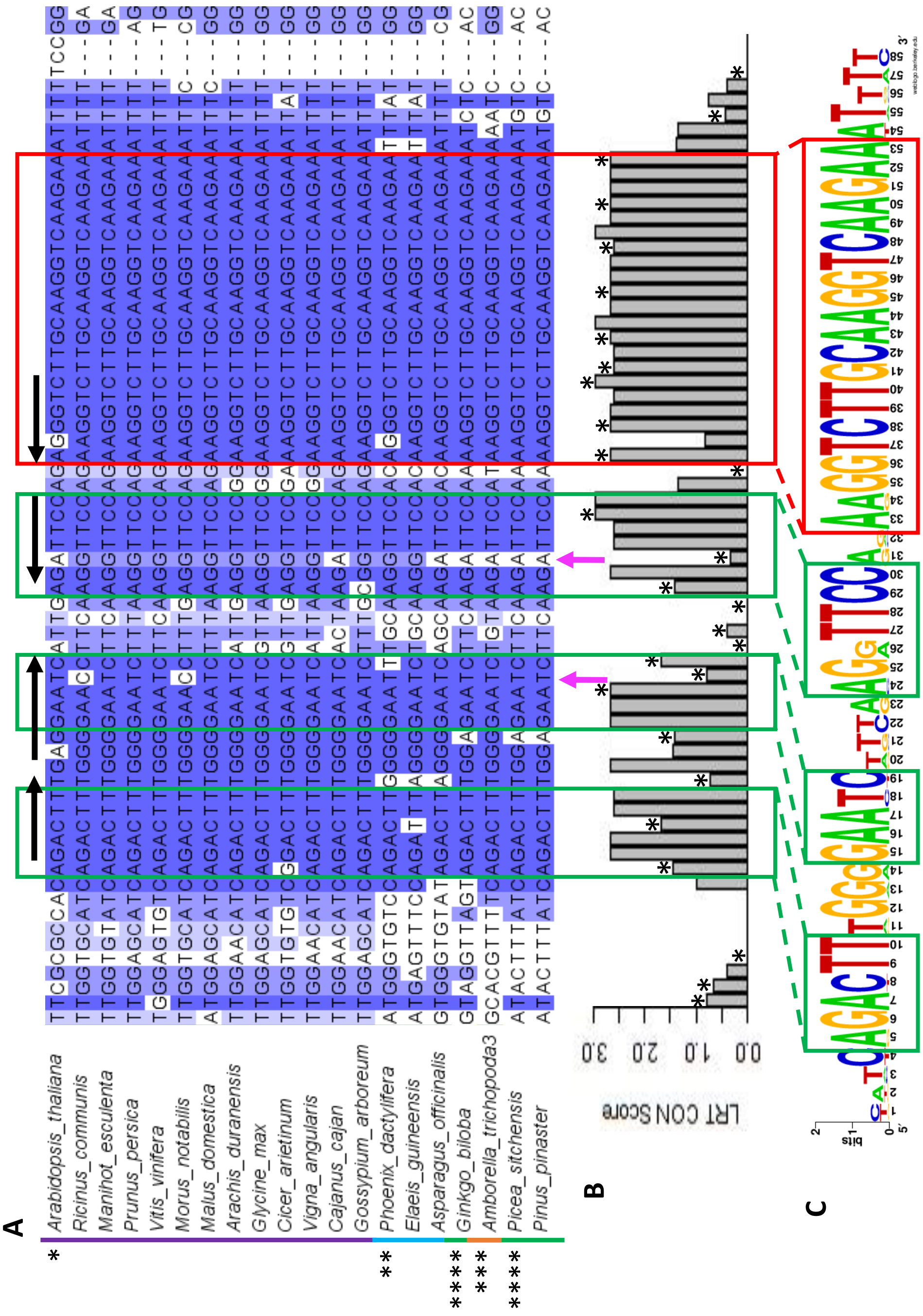


Figure 3.16. Conserved nucleotides flanking the binding site of TasiARF:*ARF2* homologues. A) MSA constructed from twenty *ARF2* homologues from various dicots and monocots. The TasiARF binding site is indicated by a red box, and the conserved flanking sequences in a green box in A to E. Black arrows above the MSA indicate base pairing from the predicted secondary structure in D and E. Asterisks indicate plant classification where '*' indicates dicots, and '**' monocots. Pink arrows indicate positions with nucleotide variations which are still compatible with base pairing and were considered conserved despite a lower phyloP score. Pink arrows indicate nucleotide variations which are still compatible with base pairing and were considered conserved despite a lower phyloP score. B) The phyloP score at each nucleotide position around the binding site. '*' denote a degenerate nucleotide site in reference to the *Arabidopsis thaliana* sequence. C) The sequence logo of the binding site and conserved flanking sequences. The height of each nucleotide indicates its relative frequency at this position. D) The secondary structure and the probability of base pairing predicted from the consensus sequence from the MSA. E) conservation annotation of the consensus sequence from the MSA. Colours represent the number of base pairs types (ie. AU, UA, CG, GC, UG, GU), and hue the number of non-conserved nucleotides at that position.

tasiARF:*ARF3* – Binding site 1 – *ARF3* contains two binding sites for tasiARF. For the 5’ binding site (binding site 1) an MSA was constructed from *ARF3* homologues from diverse lineages ranging from dicots, monocots, *Amborella trichopoda* and gymnosperms. Three conserved sequences were found upstream of the binding site across these lineages (Figure 3.17) and had sequence complementarity that was predicted to form a stem-loop that incorporated the 5’ end of the binding site (Figure 3.17 A & D). Nucleotide variations which were still compatible with base pairing were found at two positions in the conserved sequence in multiple homologues. In one of these positions one variant (G) appeared to be more common in homologues from dicot and monocot species, whereas the another (A) more common in homologues from more ancient lineages (*Amborella trichopoda* and gymnosperms).



A
B
C
Arabidopsis_thaliana
Ricinus_communis
Manihot_esculenta
Prunus_persica
Vitis_vinifera
Morus_notabilis
Malus_domestica
Arachis_duranensis
Glycine_max
Cicer_arietinum
Vigna_angularis
Cajanus_cajan
Gossypium_arboreum
Phoenix_dactylifera
Elaeis_guineensis
Asparagus_officinalis
Ginkgo_biloba
Amborella_trichopoda3
Picea_sitchensis
Pinus_pinaster
LRT CON Score
0.0
1.0
2.0
3.0
bits
5'
3'

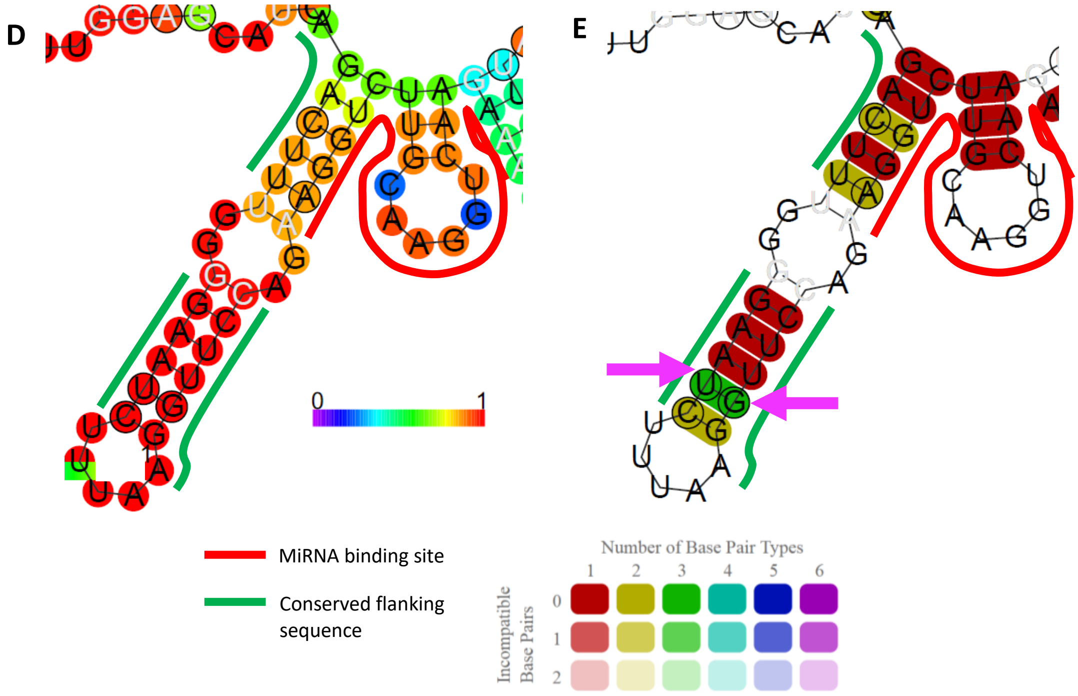
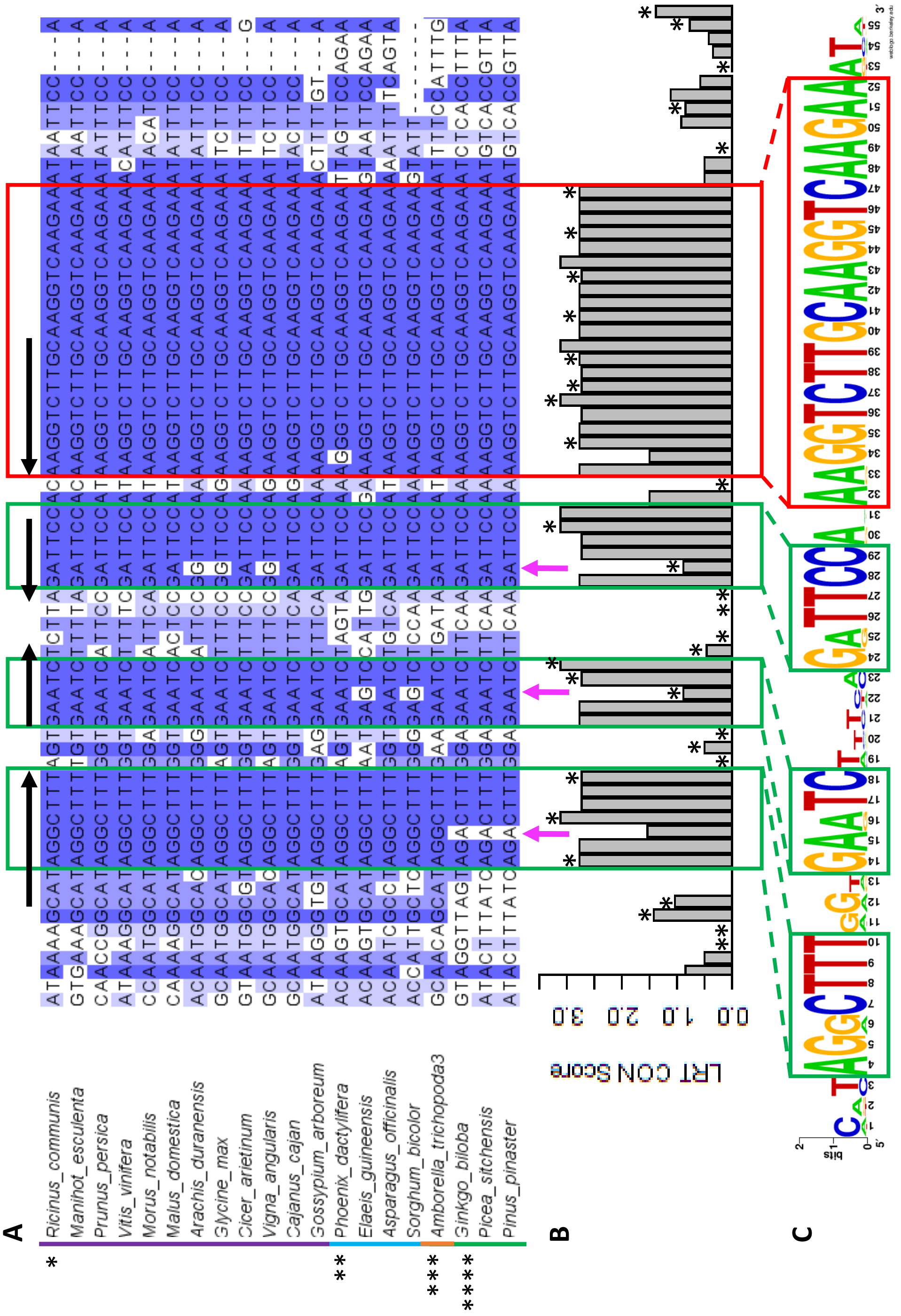


Figure 3.17. Conserved nucleotides flanking the 5' binding site of TasiARF:*ARF3* homologues (binding site 1). A) MSA constructed from twenty *ARF3* homologues from various lineages of land plants. The TasiARF 5' binding site in ARF3 (binding site 1) is indicated by a red box, and the conserved flanking sequences in a green box in A to E. Black arrows above the MSA indicate base pairing from the predicted secondary structure in D and E. Asterisks indicate plant classification where '*' indicates dicots, '**' monocots, '***' *A. trichopoda*, and '****' gymnosperms. . Pink arrows indicate positions with nucleotide variations which are still compatible with base pairing and were considered conserved despite a lower phyloP score. B) The phyloP score at each nucleotide position around the binding site. '*' denote a degenerate nucleotide site in reference to the *Arabidopsis thaliana* sequenceC) The sequence logo of the binding site and conserved flanking sequences. The height of each nucleotide indicates its relative frequency at this position. D) The secondary structure and the probability of base pairing predicted from the consensus sequence from the MSA. E) conservation annotation of the consensus sequence from the MSA. Colours represent the number of base pairs types (ie. AU, UA, CG, GC, UG, GU), and hue the number of non-conserved nucleotides at that position. The binding sites shown in the MSA correspond with the binding site 1 for all *ARF3* homologues which have two binding sites (dicots, monocots and *A. trichopoda*). The same *ARF3* homologue for gymnosperms was used in the MSA for both binding sites as they only have one binding site.

tasiARF:*ARF3* – Binding site 2 – For the 3′ tasiARF binding site in *ARF3* (binding site 2), a MSA was constructed from the same *ARF3* homologues used for binding site 1. Similar to binding site 1, three conserved sequences were found directly upstream of the binding site across all lineages (Figure 3.18) and had sequence complementarity that was predicted to form a stem-loop that incorporated the 5′ end of the tasiARF binding site (Figure 3.18 A & D). Nucleotide variations which were still compatible with base-pairing were found at two positions in the conserved sequence in multiple homologues. These conserved flanking sequences were highly similar between binding site 1 and 2, only differing at four nucleotide positions.



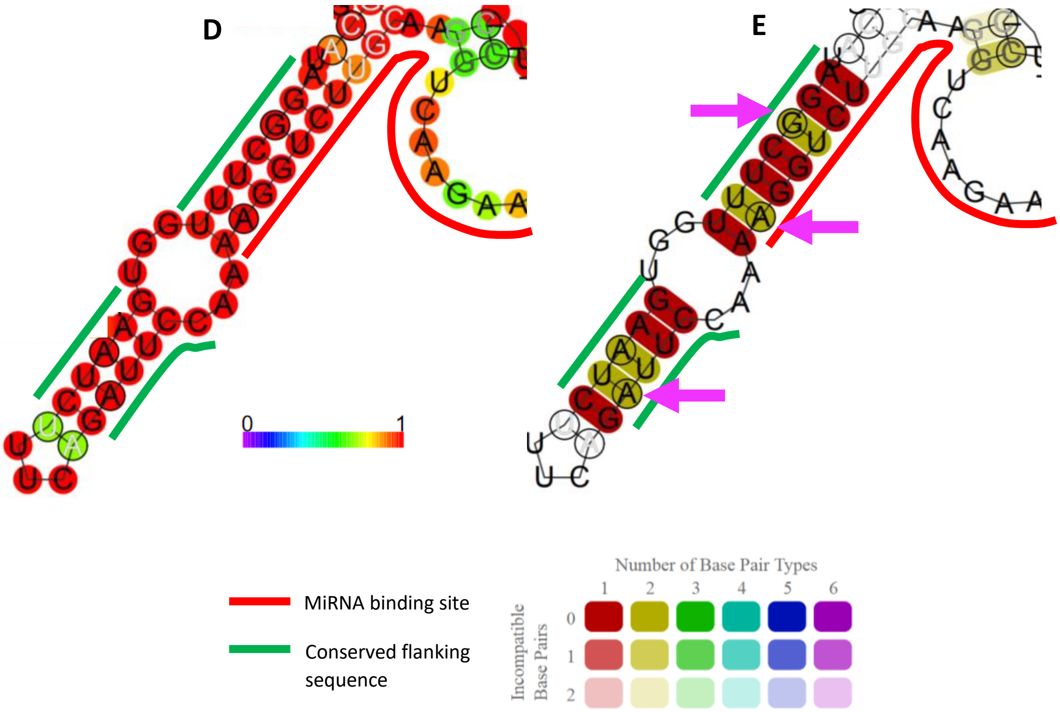
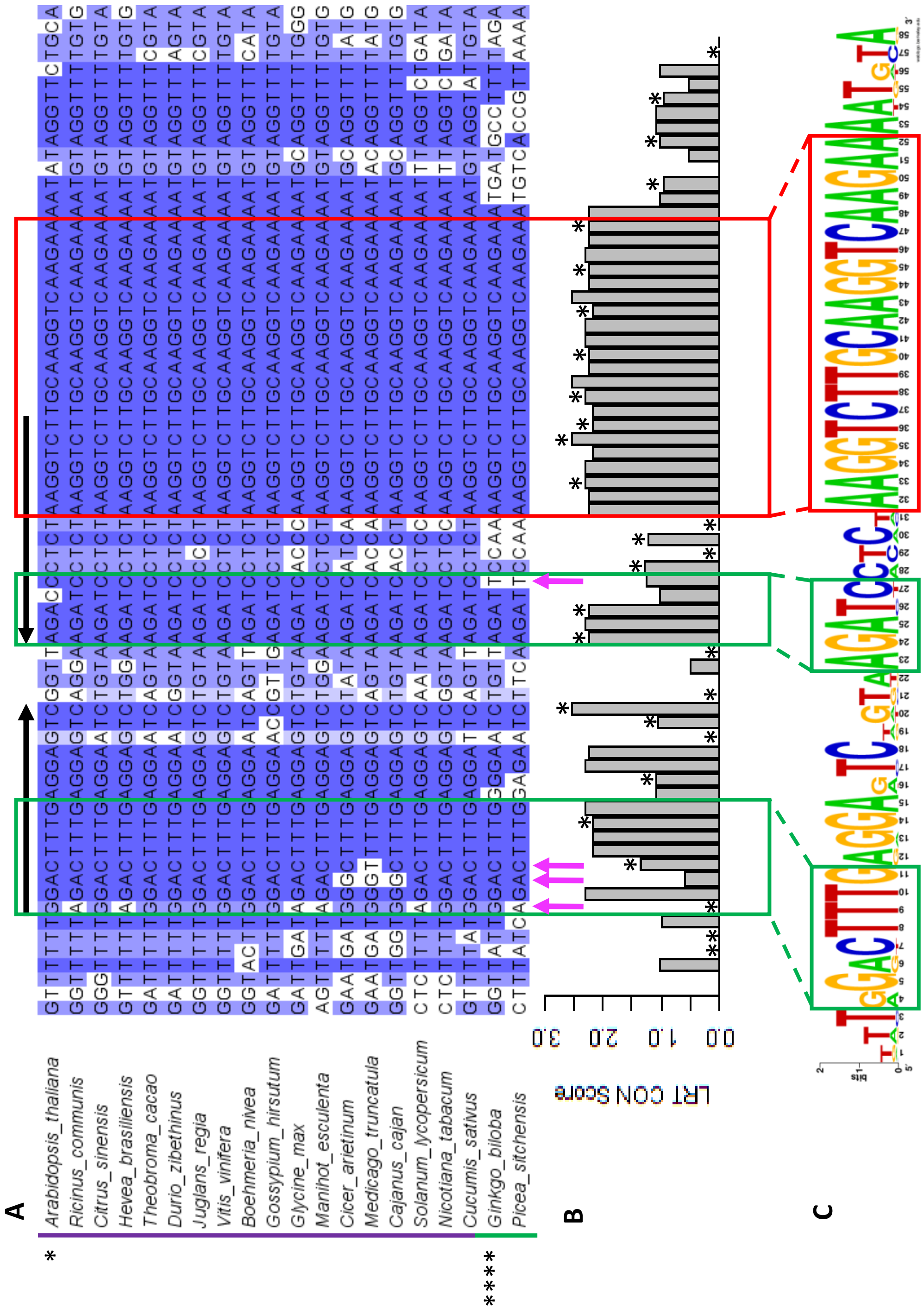


Figure 3.18. Conserved nucleotides flanking the 3’ binding site of TasiARF:*ARF3* homologues (binding site 2). A) MSA constructed from twenty *ARF3* homologues from various lineages of land plants. The TasiARF 3’ binding site in ARF3 (binding site 2) is indicated by a red box, and the conserved flanking sequences in a green box in A to E. Black arrows above the MSA indicate base pairing from the predicted secondary structure in D and E. Asterisks indicate plant classification where ‘*’ indicates dicots, ‘**’ monocots, ‘***’ *A. trichopoda*, and ‘****’ gymnosperms. Pink arrows indicate positions with nucleotide variations which are still compatible with base pairing and were considered conserved despite a lower phyloP score. B) The phyloP score at each nucleotide position around the binding site. ‘*’ denote a degenerate nucleotide site in reference to the *Arabidopsis thaliana* sequence. C) The sequence logo of the binding site and conserved flanking sequences. The height of each nucleotide indicates its relative frequency at this position. D) The secondary structure and the probability of base pairing predicted from the consensus sequence from the MSA. E) conservation annotation of the consensus sequence from the MSA. Colours represent the number of base pairs types (ie. AU, UA, CG, GC, UG, GU), and hue the number of non-conserved nucleotides at that position. The binding sites shown in the MSA correspond with the 3’ binding site for all *ARF3* homologues which have two binding sites (dicots, monocots and *A. trichopoda*). The same *ARF3* homologue for gymnosperms was used in the MSA for both binding sites as they only have one binding site.

tasiARF:*ARF4* – Binding Site 1 – Like *ARF3*, *ARF4* also contains two binding sites for tasiARF. For the 5’ binding site (binding site 1), a MSA constructed of *ARF4* homologues from eighteen dicot and two gymnosperm species showed two conserved sequences upstream of the binding site (Figure 3.19). The 5’ conserved sequence was in-part complementary to the tasiARF binding site with which it was predicted to form an RNA stem-loop (Figure 3.19 D). Nucleotide variations which were still compatible with base-pairing were found at four positions in the conserved sequence in multiple homologues (Figure 3.19 E).



s

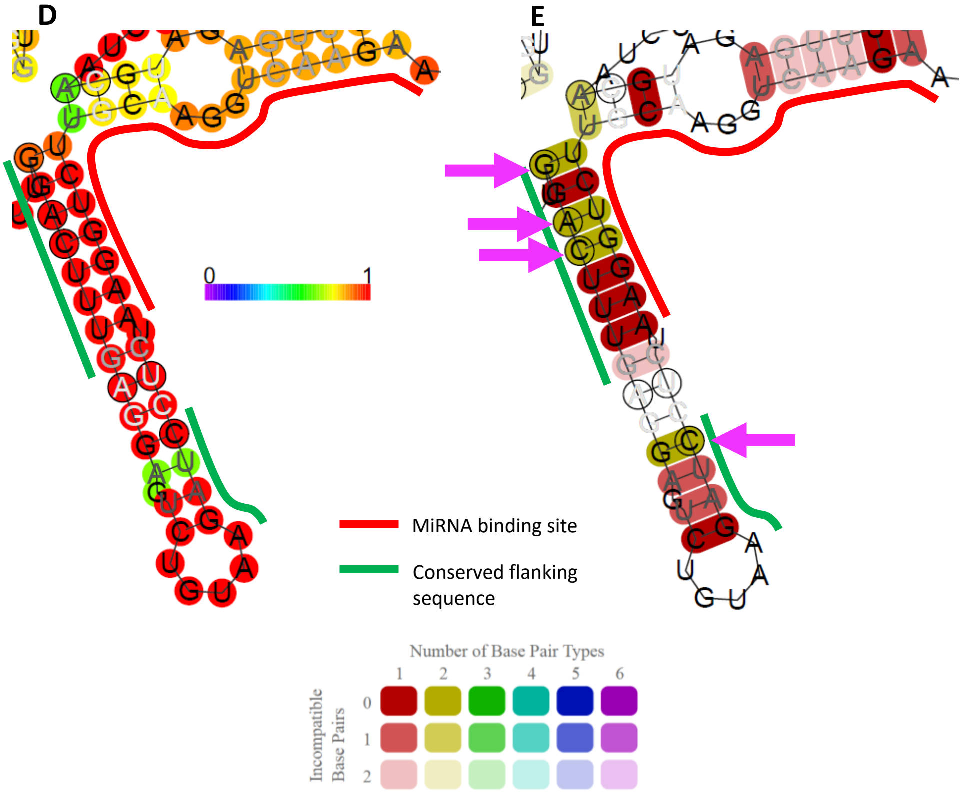
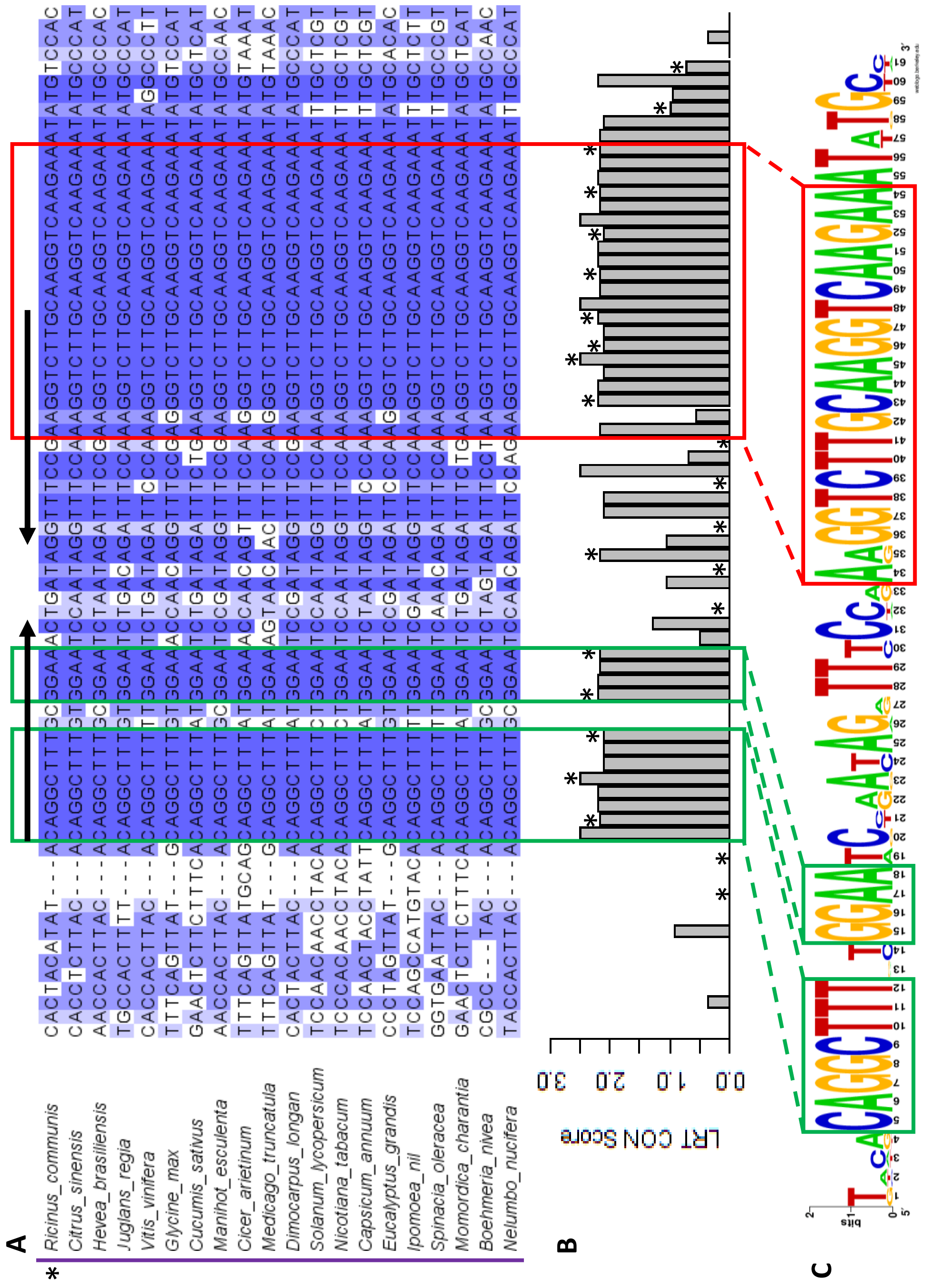


Figure 3.19. Conserved nucleotides flanking the 5’ binding site of TasiARF:*ARF4* homologues. A) MSA constructed from twenty ARF4 homologues from various lineages of land plants. The TasiARF 5’ binding site in ARF4 (binding site 1) is indicated by a red box, and the conserved flanking sequences in a green box in A to E. Black arrows above the MSA indicate base pairing from the predicted secondary structure in D and E. Asterisks indicate plant classification where ‘*’ indicates dicots, and ‘****’ gymnosperms. Pink arrows indicate positions with nucleotide variations which are still compatible with base pairing and were considered conserved despite a lower phyloP score. B) The phyloP score at each nucleotide position around the binding site. ‘*’ denote a degenerate nucleotide site in reference to the *Arabidopsis thaliana* sequence. C) The sequence logo of the binding site and conserved flanking sequences. The height of each nucleotide indicates its relative frequency at this position. D) The secondary structure and the probability of base pairing predicted from the consensus sequence from the MSA. E) conservation annotation of the consensus sequence from the MSA. Colours represent the number of base pairs types (ie. AU, UA, CG, GC, UG, GU), and hue the number of non-conserved nucleotides at that position. The binding sites shown in the MSA correspond with the 5’ binding site for all *ARF4* homologues which have two binding sites (dicots, monocots and *A. trichopoda*). The same *ARF4* homologue for gymnosperms was used in the MSA for both binding sites as they only have one binding site.

tasiARF:*ARF4* – Binding Site 2 – For the *ARF4* binding site (binding site 2), a MSA was constructed using twenty homologues from diverse dicot species (Figure 3.20). Conserved sequences upstream of the tasiARF binding site were identified which had complementarity to the binding site with which it was predicted to form a stem-loop (Figure 3.20 D).



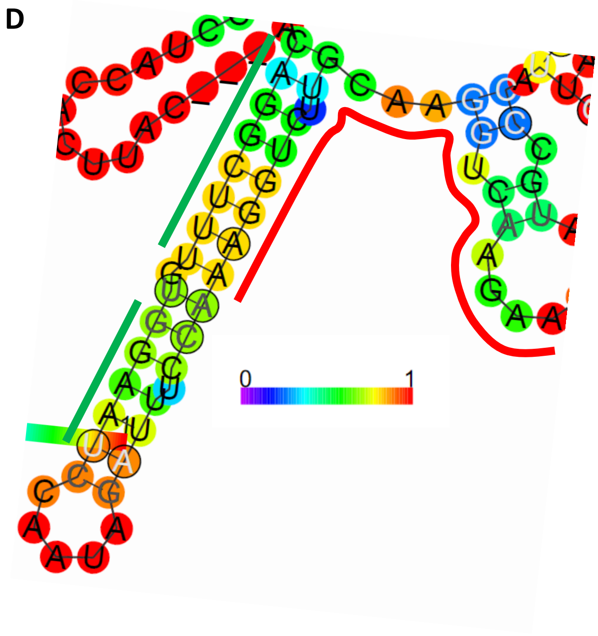


Figure 3.20. Conserved nucleotides flanking the 3’ binding site of TasiARF:*ARF4* homologues. A) MSA constructed from twenty *ARF4* homologues from various dicots. The TasiARF 3’ binding site in ARF4 (binding site 2) is indicated by a red box, and the conserved flanking sequences in a green box in A to D. Black arrows above the MSA indicate base pairing from the predicted secondary structure in D. Asterisks indicate plant classification where ‘*’ indicates dicots. B) The phyloP score at each nucleotide position around the binding site. ‘*’ denote a degenerate nucleotide site in reference to the Arabidopsis thaliana sequence. C) The sequence logo of the binding site and conserved flanking sequences. The height of each nucleotide indicates its relative frequency at this postion. D) The secondary structure and the probability of base pairing predicted from the consensus sequence from the MSA. The binding sites shown in the MSA correspond with the 3’ binding site for all *ARF4* homologues which have two binding sites (dicots, monocots and *A. trichopoda*). The same *ARF4* homologue for gymnosperms was used in the MSA for both binding sites as they only have one binding site.

3.2.7 Evidence for RNA secondary structure formation for the tasiARF:*ARF* conserved flanking sequences

When comparing the consensus around the binding sites of *ARF2*, *ARF3* and *ARF4*, commonalities were found in the conserved flanking sequences. A high degree of sequence identity between the conserved flanking sequences is found despite considerable variation elsewhere at this location (Figure 3.21A). Furthermore, any nucleotide variations between the conserved sequences were either compensatory substitution or single nucleotide substitutions (‘U’ to ‘C’, ‘A’ to ‘G’, and *vice versa*) which were still compatible with base-pairing to form a stem-loop (Figure 3.21). This suggests that over evolutionary time, despite neutral drift, a similar RNA secondary structure is being selected for. Further supporting this is that the least amount of conservation is found in the nucleotide positions corresponding to the loop region of these stem-loops. Therefore, this suggests that there is strong selection for this stem-loop RNA secondary structure next to the tasiARF binding site in the *ARF* family members which may suggest a functional role in tasiRNA-mediated regulation.

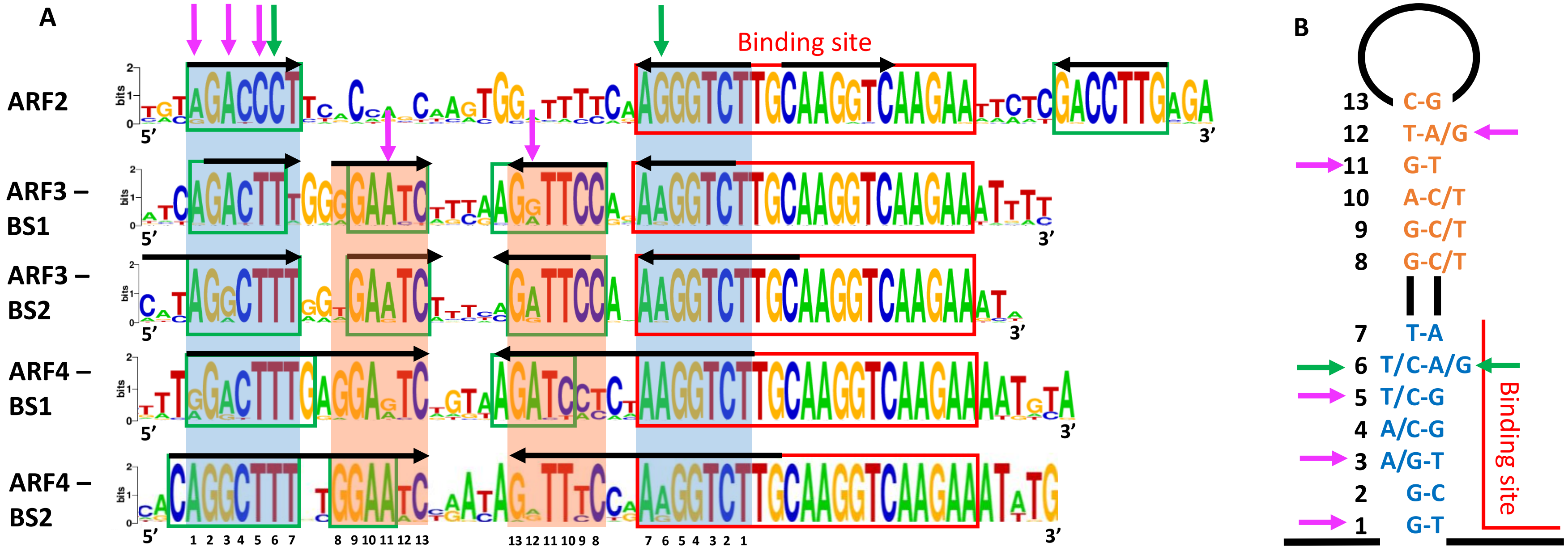
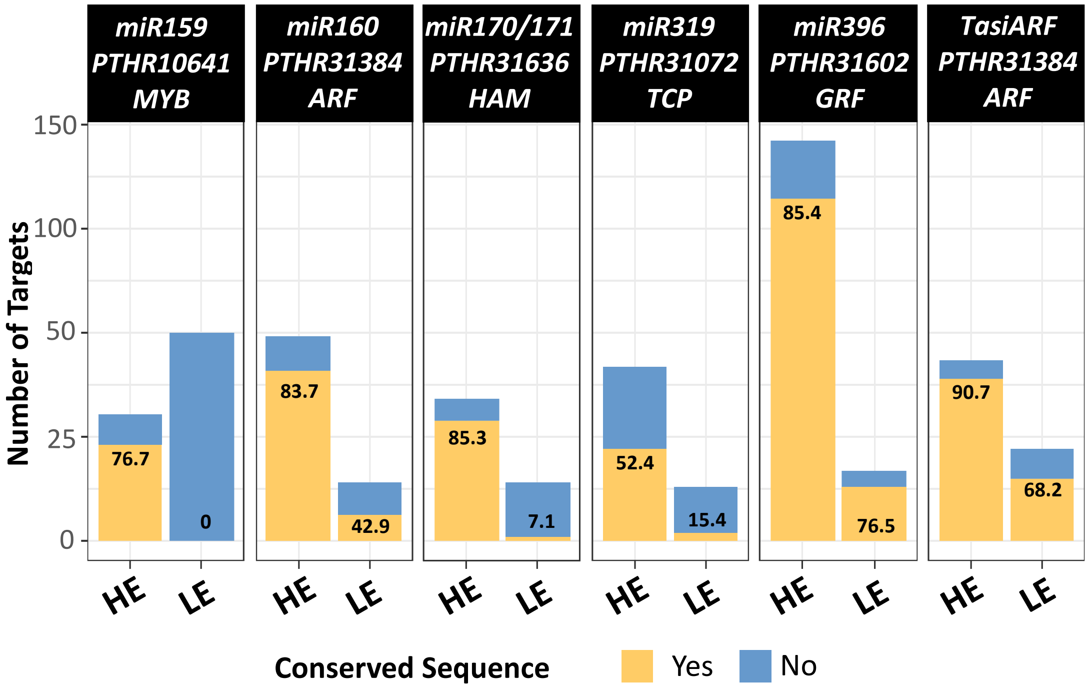


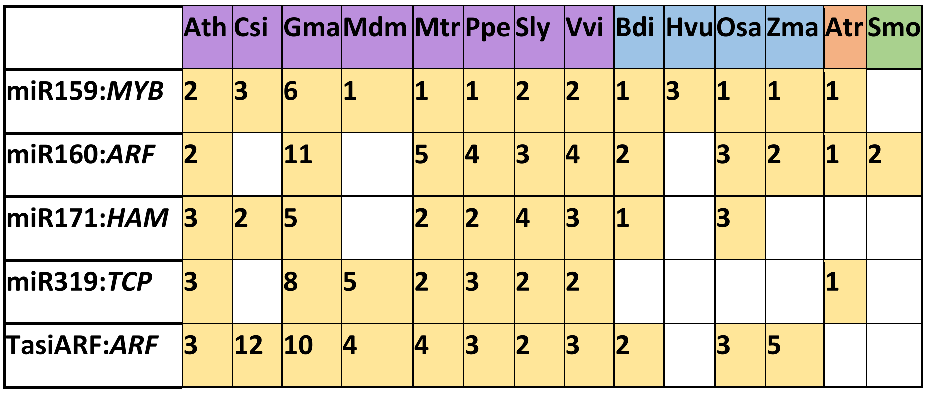
Figure 3.21. Comparison of conserved sequences flanking the binding site in TasiARF targets. A) A comparison of the consensus sequences flanking the binding sites of *ARF2*/*ARF3*/*ARF4* (note the *ARF3* and *ARF4* have two binding sites each). The binding sites are indicated by the red box and the conserved flanking sequences by the green box. Black arrows indicate base pairing from the predicted secondary structure. Nucleotides shaded in blue and orange correspond to the nucleotides coloured in Figure 3.21B. B) Nucleotide variations in the conserved flanking sequences which are still compatible for stem loop formation. Positions with single nucleotide substitutions which are still compatible for base pairing are indicated by pink arrows ('U' to 'C', 'A' to 'G', and *vice versa*). The green arrows indicate compensatory substitutions where there are variations at both nucleotide positions and are still compatible with base pairing. Note that nucleotide 'U' is shown as 'T'. Numbers in black correspond to the nucleotide positions of the stem in Figure 3.1 A and B.

3.2.8 Conserved sequences flanking the miRNA binding site are enriched in HE targets

Having identified conserved sequences flanking the binding site in multiple target families, it was investigated how these sequences were distributed among HE and LE target homologues (Figure S3, Figure S4). For all target families, a higher percentage of HE targets possessed the conserved sequences compared to LE targets indicating that these sequences are enriched in targets subjected to strong miRNA-mediated regulation (Figure 3.22 A). However, the miR159-mediated regulated MYB targets still demonstrated the strongest and most striking enrichment. Compared to the other miRNA-target pairs, no LE targets were identified with the conserved sequences.

For all miRNA-target family modules, conserved sequences flanking the binding site were identified in HE targets from species spanning beyond dicots (Figure 3.22 B). Furthermore, for the miR159:*MYB* and miR319:*TCP* family module conservation was found to extend to *Amborella trichopoda*, and, strikingly, even to the distantly related lycophyte, *Selaginella moellendorffii*, in the miR160:*ARF* family module. Therefore, these results further support a functional importance of these sequences in miRNA-mediated regulation that is deeply conserved.





	Ath	Csi	Gma	Mdm	Mtr	Ppe	Sly	Vvi	Bdi	Hvu	Osa	Zma	Atr	Smo
miR159:*MYB*	2	3	6	1	1	1	2	2	1	3	1	1	1	
miR160:*ARF*	2		11		5	4	3	4	2		3	2	1	2
miR171:*HAM*	3	2	5		2	2	4	3	1		3			
miR319:*TCP*	3		8	5	2	3	2	2					1	
TasiARF:*ARF*	3	12	10	4	4	3	2	3	2		3	5		

Figure 3.22. Presence of conserved sequences flanking the binding site in HE and LE targets in multiple miRNA-target family modules. A) Analysis for the conserved sequences in the HE and LE targets of the primary target family. Genes with conserved sequences are indicated in yellow and genes without conserved sequences are indicated in blue. Numbers above bars indicate the total number of genes analysed. The number of genes with the conserved sequence out of total genes is expressed as a percentage in the yellow bars. Note that the graph for miR159:MYB is the same as Figure 3.7 and is included for comparison. B) The number of HE targets possessing the conserved sequences flanking the binding site across species. Dicots are highlighted in purple, monocots in blue, *A. trichopoda* in orange and *S. moellendorffii* in green. Conserved sequences from family members of the same miRNA:Target family module were analysed together [ie. miR160:ARF10 & ARF17; miR319:TCP2 & TCP4; TasiARF:ARF2, ARF3 & ARF4].

3.2.9 Mutations to the conserved flanking sequences in *ARF10* impacts miR160-mediated regulation

To further investigate if these conserved sequences are involved in miRNA-mediated regulation, analysis using functional genetic approaches *in planta* are required. The miR160 target, *ARF10* was chosen due to its highly conserved flanking sequences which are predicted to form an RNA secondary structure with high confidence. Additionally, plants where miR160-mediated regulation of *ARF10* has been perturbed are well characterized and have an easily distinguishable phenotype (Liu et al., 2007). Synonymous mutations were introduced into the conserved sequences flanking the miR160 binding site. These consisted of two single nucleotide mutations in the 5’ conserved sequence and three mutations in the 3’ conserved sequence which did not change the amino acid sequence (*ARF10-FM*) (Figure 3.23 A). As the RNA secondary structure may impact miRNA-mediated regulation, mutations were chosen that altered the predicted RNA secondary structure (Figure 3.23 B). To determine if these mutations would impact miR160-mediated regulation of *ARF10*, *ARF10-FM* was compared to an *ARF10* construct without mutations in the conserved sequence (*ARF10-WT*) and a construct with mutations to the miR160 binding site rendering it resistant to miR160-mediated regulation (*rmARF10*) (Figure 3.23 A). All *ARF10* variants were fused to a CaMV 2x35s promoter for wide and constitutive expression as miR160 is widely expressed across tissues (Mallory et al., 2005). Therefore, any dysregulation to miR160-mediated regulation of the *ARF10* variants will be easily identifiable.

All *ARF10* variant constructs were individually transformed into Arabidopsis. Primary transformants for each variant was then phenotyped for a mutant leaf curl phenotype which has been previously reported in transgenic plants overexpressing miR160-resistant *ARF10* (Liu et al., 2007). Phenotypic defects were catergorised by severity into ‘No leaf curl’, where plants displayed no leaf curl and were indistinguishable from wild type plants; ‘Weak’, where plants displayed some leaf curl with up to one leaf curled with the abaxial side visible from an aerial view; and ‘Strong’, where plants displayed two or more leaves curled with the abaxial side visible (Figure 3.24 A). Although Liu et al. (2007) previously reported serrated leaves as a morphological defect, only one primary transformant displayed this phenotype. Results found a majority of *ARF10-WT* primary transformants displayed a ‘No leaf curl’ phenotype (73%) whereas this was just over half for *ARF10-FM* (52%). *ARF10-WT* primary transformants also demonstrated less severe phenotypes with only four plants (3%) categorized as having a ‘Strong’ mutant phenotype compared to *ARF10-FM* (11%). *rmARF10* displayed the least primary transformants without a mutant phenotype (39%) and the most plants with a ‘Strong’ mutant phenotype (21%), although these numbers were not statistically different from *ARF10-FM*. However, lethality is a previously reported defect for miR160-resistant *ARF10*. Therefore, given the difficulty recovering *rmARF10* seedlings and that many plants died before phenotyping, it is likely that the strength of dysregulation in *rmARF10* was underestimated. Nevertheless, the conserved sequences flanking the binding site appear to influence miR160-mediated regulation of *ARF10*.

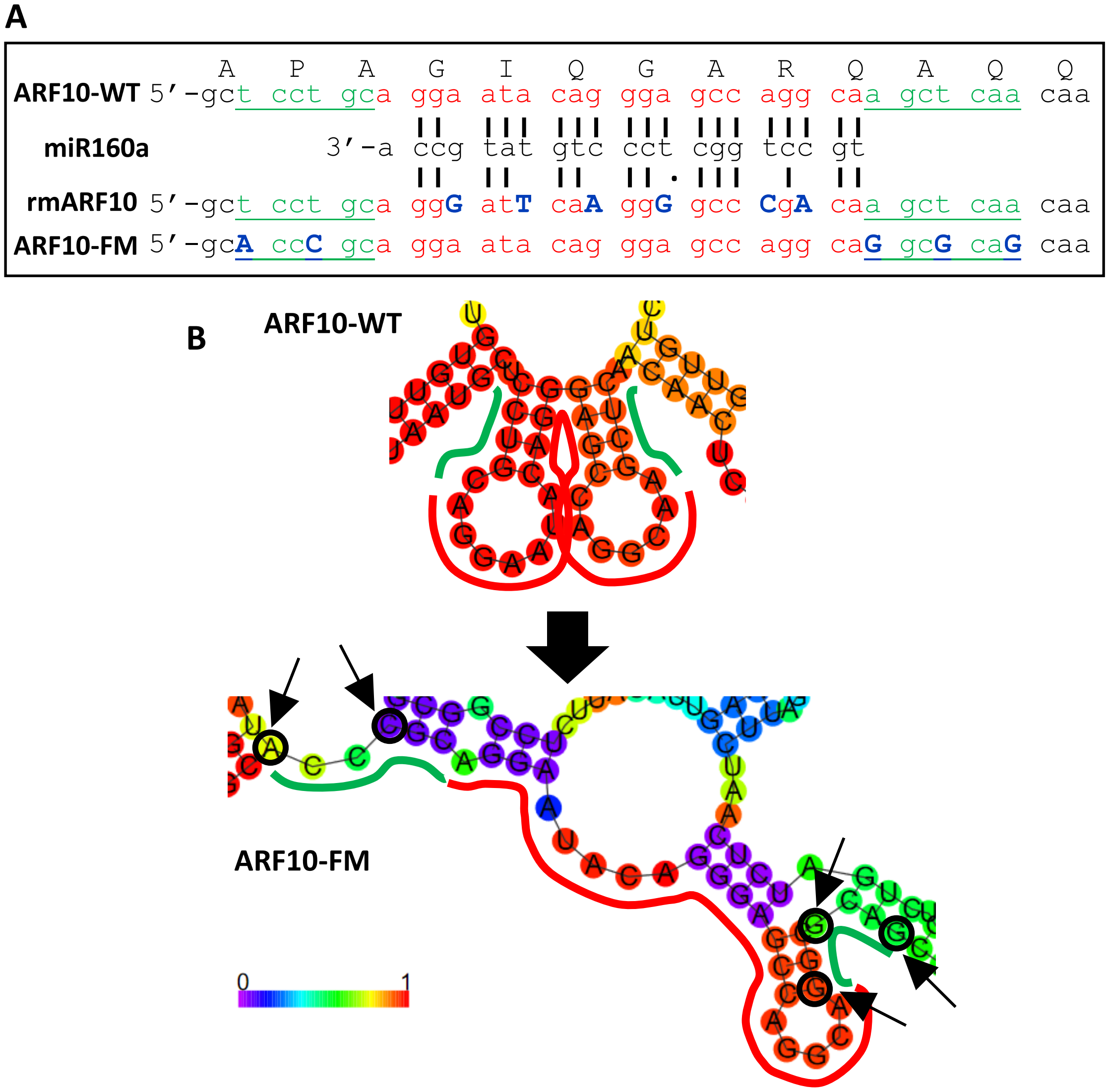


Figure 3.23. ARF10 Variant transgenes constructs. A) Alignments of the miR160a and ARF10 transgenic constructs used to transform Arabidopsis. The binding site is indicated in red, and the conserved flanking sequences in green. For *rmARF10* and *ARF10-FM*, mutated nucleotides in the binding site and conserved flanking sequences are capitalised and highlighted in blue. No amino acids were changed between transgenic constructs. B) Predicted secondary structure showing the base pair probability of ARF10-WT and ARF10-FM. The binding site is indicated by the red line, and conserved sequence in green. Nucleotides changed in ARF10-FM are circled and indicated by arrows. Both sequences are from Arabidopsis.

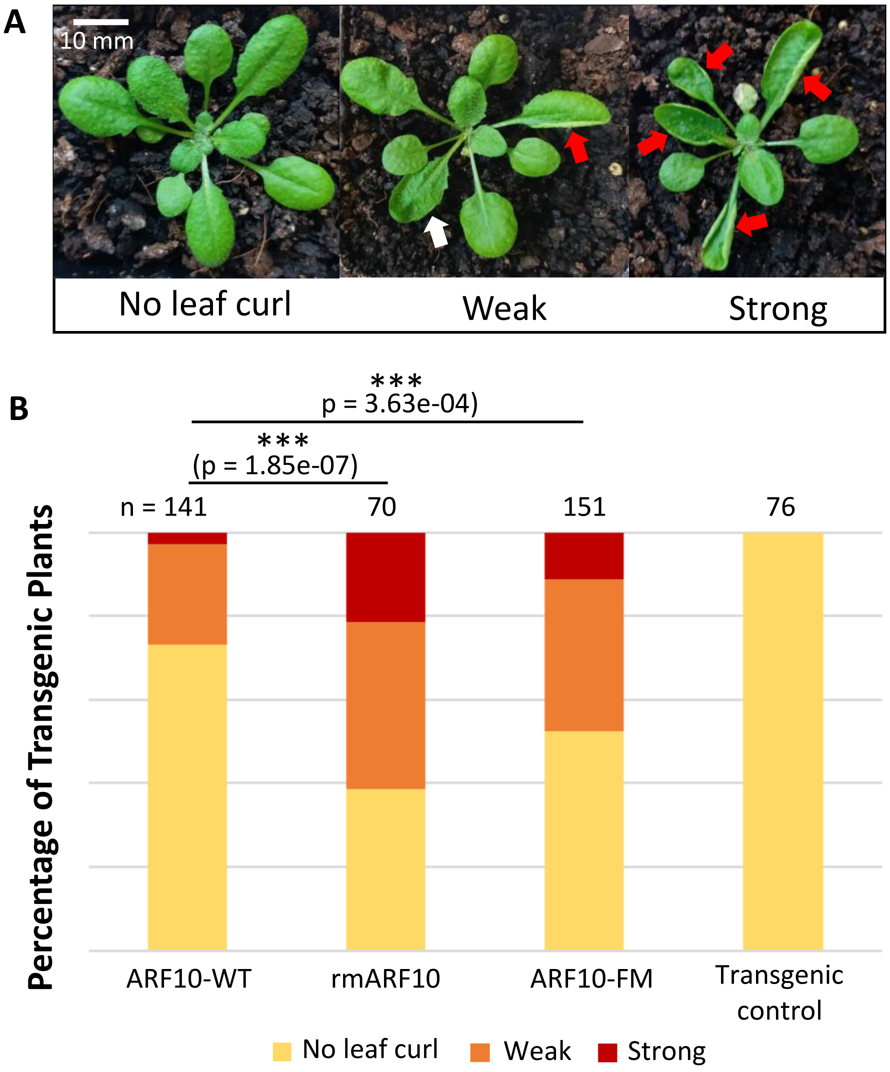


Figure 3.24. *ARF10* transgenic plants showing leaf curl mutant phenotype. Phenotypes of four-week-old Arabidopsis plants showing mutant rosette phenotypes. A) The red arrows indicate leaves curled so that the abaxial side is showing when viewed from the top. White arrows indicate leaves with apparent leaf curl but without the abaxial side is showing. Phenotype scoring was divided into three categories. ‘No leaf curl’ indicates plant exhibited no presence of leaf curl; ‘Weak’ indicate plants exhibited presence of leaf curl with up to one leaf curled with the abaxial side visible from an aerial view; ‘Strong’ indicates plant exhibited two or more leaves curled with the abaxial side visible. White arrows indicate leaves with the presence of leaf curl. B) Percentage of ARF10 primary transformants showing a ‘No leaf curl’, ‘Weak’ and ‘Strong’ phenotype. ‘n’ indicates the number of plants analysed. A vector only transgenic control construct was also included.

3.3 Discussion

Similar to previous studies, results from this Chapter show that using miRNA-target binding site complementarity is insufficient to predict a physiologically relevant MTI (Brousse et al., 2014; Liu et al., 2014; Zheng et al., 2017). Therefore, in support of previous findings, this suggests that there exists factors additional to miRNA-binding site complementarity for miRNA-mediated regulation in plants (Gu et al., 2012; Li et al., 2012; Zheng et al., 2017; Yang et al., 2020). Contributing to these studies, TRUEE was used to analyse multiple highly conserved miRNAs where, given the evolutionary history, it would be feasible for such factors to arise. Firstly, these results found that, for each miRNA, a single highly conserved gene family predominated its MTIs and that having additional gene families was rare. Secondly, conserved sequences flanking the binding site were found to correlate with genes with high experimental evidence as targets. In many instances, these flanking sequences were predicted to form RNA secondary structures with the miRNA-binding sites. Lastly, we demonstrated that the conserved sequence flanking the miR160 binding site in *ARF10* influenced the miR160-mediated regulation of *ARF10*.

3.3.1 TRUEE analysis demonstrates conserved MTIs predominate across species

Despite the myriad of predicted targets, only a small number of these predicted targets have been experimentally validated to date. An explanation is that regulatory constraints exists which limit the scope of miRNA-mediated regulation on the plant transcriptome (Li et al., 2014).

Supporting this, TRUEE analysis demonstrated experimentally that conserved MTIs predominate across species, with only a few conserved secondary target families being identified (Figure 3.2; Table 3.1). As most of the conserved targets are regulatory genes core for plant biology, their post-transcriptional regulation by miRNA appears indispensable and under strong selective pressure (Reviewed in Samad et al., 2017). Given the miRISC is an independent regulatory unit, the expression of these conserved miRISCs (i.e. the temporal and spatial expression level of the miRNA) will be under strong selection of the regulatory requirements of the function of the conserved targets, and therefore will constrain the acquisition of additional targets. Such is the specificity, for some miRNAs, TRUEE analysis identified HE targets that were almost exclusively one gene family (eg. miR160:*ARF*; miR166:*HD-ZIPIII*; miR172:*AP2*) (Figure 3.2). No conserved miRNA family was found to have switched primary target families which is consistent with a lack of examples from the literature. To date, the only example is miR827, where its primary target family appears to have transitioned from *PLASMA-MEMBRANE-LOCALIZED PHOSPHATE TRANSPORTER 5* (*PHT5*) found in many angiosperms to *NITROGEN LIMITATION ADAPTATION* (*NLA*) in the *Brassicaceae* and *Cleomaceae* (Lin et al., 2018). Such a rare scenario in plants contrasts to that in animals, where it is not uncommon for a single miRNA family to target a large number of distinct target families (Kedde et al., 2007; Kedde et al., 2010; Lustig et al., 2014; Humphries & Yang, 2015; Cheng et al., 2017; Magenta et al., 2017; Iwai et al., 2018; Vahabi et al., 2019; Zhou et al., 2019; Tian et al., 2020). Potentially underlying these differences are the high complementarity requirement of MTIs in plants, and the strength of the silencing outcomes, both of which seem much higher in plants than for MTIs in animals.

3.3.2 Multiple target families of a conserved miRNA are likely to be functionally related

It was hypothesized that a consequence of this regulatory constraint is that additional acquired targets must have a MTI that is compatible to regulatory conditions defined by the primary MTI (Li et al., 2014). This is because the expression pattern of the miRNA will be dictated by the desired regulatory outcome of the primary target family, and so the regulation of any additional targets must be achieved in the context of this miRNA expression pattern.

Supporting this scenario, TRUEE analysis found many of the secondary target families identified had MTIs which were functionally related to the primary target family (Figure 3.2 & Table 3.1). For miR395, its expression is induced under sulphate deficiency which leads to the regulation of its primary target family, *APS* (Liang et al., 2010). To acquire and maintain an additional target, the desired regulatory outcome of the secondary target family, *SULTR*, must also be downregulation under these same conditions. Similarly, for miR398, miRNA expression is induced under oxidative stress for the downregulation of its primary and secondary target families, *SOD* and *COX*, respectively (Sunkar et al., 2006; Yamasaki et al., 2007). Furthermore, *LAC* and *PLANTACYANIN*, which were also found as a secondary target family for miR398 in analysis, are also copper proteins like *SOD* (Abdel-Ghany & Pilon, 2008). These gene families are more often reported to be the main targets of miR397 and miR408, respectively, which, like miR398, are also miRNAs induced by copper deficiency (Pilon, 2017). Additionally, the copper transporter, *COPT*, was also identified as a secondary target family of miR398 and is also involved in the same copper pathway as *SOD* (Pilon, 2017). *COPT* has not previously been reported as a target family and may suggest a new MTI for miR398 in some dicot species. A study by Naya et al. (2014) in *Phaseolus vulgaris* found lower expression of *COPT* in plants overexpressing miR398. However, this difference was nowhere near as great compared to the *SOD* family member, *COPPER SUPEROXIDE DISMUTASE* (*CSD1*), where mRNA levels was severely reduced in miR398 overexpression plants compared to the control.

Examples where the primary and secondary target family is functionally related was also found for MTIs reported in literature. The miR399 targets, *PHOSPHATE2* (*PHO2*), a ubiquitin conjugase protein, and *IPS1*, a non-coding RNA, play distinct roles in phosphate (Pi) deficiency. *PHO2* acts as a Pi transporter, and *IPS1* acts as a miRNA mimic which sequesters miR399 to fine-tune *PHO2* activity and the Pi deficiency response (Franco-Zorrilla et al., 2007). Similarly, for miR827, which appeared to have changed targets from *PHT5* to *NLA* in *Brassicaceae* and *Cleomaceae* (mentioned above), many similarities exist between its two target families. Both *NLA* and *PHT5* encode proteins with an *SYG1/PHO81/XPR1* (*SPX*) domain and are involved in Pi deficiency where they function in Pi transport and Pi storage, respectively. Furthermore, the conditions at which miR827 is expressed (under Pi deficiency) is still the same across lineages (Lin et al., 2018). It may be that such a change in target was only permissible because *NLA* is still regulated under the same conditions as *PHT5* which would be necessary as at one point both genes would be targets of miR827 simultaneously.

TRUEE analysis also found that the secondary target families were less conserved and had fewer members which has also been reported in literature (Table 3.1). For example, in addition to the *GRF* primary target family, miR396 also targets a basic Helix-Loop-Helix, *bHLH74*, where both are involved in leaf development (Debernardi et al., 2012). Whereas *GRF* was found to be conserved across dicots, monocots and *Amborella trichopoda*, *bHLH74* was only found in the *Brassicaceae* and *Cleomaceae* (Debernardi et al., 2012). Chorostecki et al. (2012) also bioinformatically predicted and experimentally validated several other primary and secondary target pairs which were also functionally related and conserved in a narrower group of species [miR167: *ARF* & *IAA-ALANINE RESISTANT 3* (*IAR3*); miR396:*GRF,* miR396:*MMG4.7* and miR396:*FLUORESCENT IN BLUE LIGHT* (*FLU*)]. This may reflect that acquiring beneficial additional MTIs which are functionally related to the primary target family is rare and may be easily lost and further lends evidence to the predominance of the primary MTI.

3.3.3 Conserved complementarity varies greatly between miRNA-target pairs

TRUEE analysis clearly demonstrates that considering complementarity as a sole factor is insufficient in predicting an HE target across miRNA families. It is clear that complementarity requirements varied greatly between each miRNA-target family pair, with the average Expectation Score of HE targets varying from 0.4 for miR160, to 4.3 for miR398 (Figure 3.4). This implies complementarity cannot be used as a clear indicator of an HE target across miRNA families.

Consistent with this is that Liu et al. (2014) found that binding sites engineered with perfect complementarity to the miRNAs are not the most strongly silenced. Similarly, artificial miRNAs engineered with similarly high complementarity to their intended targets also varied in regulation (Deveson et al., 2013; Li et al., 2013). Some miRNA targets with suboptimal complementarities have been experimentally validated, whilst there are other predicted targets with a higher degree of complementarity for which little or no evidence has been found (Table 2.1) (Debernardi et al., 2012; Brousse et al., 2014). Therefore, this implies that additional factors are involved in the miRNA-mediated regulation of a target.

3.3.4 A role for RNA secondary structure in facilitating miRNA-mediated regulation?

Currently, the only demonstration of factors additional to complementarity in miRNA-mediated regulation are the conserved flanking sequences associated with the miR159-binding site of *MYB33* that form a predicted RNA secondary structure and which were functionally demonstrated to facilitate miR159-mediated regulation (Zheng et al., 2017). In this Chapter, further miRNA-target families with conserved flanking sequences have been identified. Given these conserved sequences are enriched in HE targets, this suggests they may be facilitating strong MTIs.

Many of these conserved sequences were also predicted to form RNA secondary structures [*ARF10*, *ARF17*, *HAM1*, *IPS1*, *ARF2*, *ARF3* and *ARF4*]. For instance, a similar RNA secondary structure was present in the *ARF2, ARF3* and *ARF4* homologues despite sequence divergence (Figure 3.16 – 3.20 D, Figure 3.21). The occurrence of nucleotide variations and compensatory substitutions that maintained base-pairing with the predicted RNA secondary structures of these *ARF* genes suggests it is these RNA secondary structures that are being selected for.

Curiously, many of the conserved flanking sequences identified were predicted to base-pair with the miRNA-binding site. Although in the first instance a highly structured miRNA-binding site may seem counter-intuitive, as accessibility may attenuate regulation, an *in vivo* assessment of RNA structure of miRNA-binding sites found them to be highly structured (Yang et al., 2020).

Additionally, similar to the RNA stem-loops associated with the miR159-binding site of *MYB33/65* (Zheng et al., 2017), the majority of conserved flanking sequences were located upstream and were predicted to base-pair with the 5’ end of the miRNA-binding site, leaving the 3’ end of the miRNA-binding site unbound (miR160:*ARF17*; miR171:*HAM1*; tasiARF:*ARF3*; tasiARF:*ARF4*). The 3’ end region of the binding site corresponds to the nucleotide positions most important for miRNA-mediated regulation (5’ end of the miRNA), as multiple studies have found that mismatches within this region preferentially attenuate regulation (Mallory et al., 2004; Schwab et al., 2005; Liu et al., 2014). Therefore, it is tempting to speculate that these structures are designed to promote accessibility to this region of the miRNA-binding site. For the miR160:*ARF10* and tasiARF:*ARF2* modules, conserved sequences were also found downstream of the binding site and were predicted to base-pair with the 3' region of the miRNA-binding site. Interestingly, in these cases, some of the 3' nucleotides of the binding site still coincided with the unpaired loop region of the stem-loop which may be leading to an open conformation and greater accessibility.

Alternatively, these base-pairings may be inhibiting accessibility to the target binding site. An *in vivo* study on the mRNA structurome in Arabidopsis also found the miRNA binding site to be more structured (Yang et al., 2020). However, they concluded that this rendered the miRNA-binding sites less accessible to miRISC prior to target cleavage. Rather, the unfolding of this secondary structure functions as a rate-limiting factor of miRISC cleavage efficiency. Only the 2 nt downstream of the miRNA-binding site were required to be unstructured for efficient target cleavage by AGO but not binding (Yang et al., 2020). This is also consistent with our results in that most secondary structures appeared upstream but not downstream of the binding site. Alternatively, these predicted structures may play a role in ribosome stalling as RNA secondary structures have been reported to cause ribosome stalling in plants (Gawronski et al., 2018). Therefore, it may be that these structures cause the ribosome to stall and delay the completion of translation which therefore increases miRNA-binding site accessibility for increased silencing. Clearly, more work is needed here to determine whether these predicted RNA secondary structures exist *in vivo* and their function, if any, on miRNA-mediated regulation.

In this Chapter, in addition to *MYB33*, the conserved flanking sequences in a second miRNA target, *ARF10*, was functionally demonstrated to be involved in miRNA-mediated regulation (Figure 3.24). Like in *MYB33*, the conserved flanking sequences also form a predicted RNA secondary structure. Furthermore, mutations to the flanking sequences, which was also predicted to alter the RNA secondary structure, attenuated miR160-mediated regulation. Thus, this further supports a role for RNA secondary structures flanking the miRNA binding site in the miRNA-mediated regulation of some targets. Other targets were found to have conserved sequences flanking the miRNA binding site which correlated with a HE target (Figure 3.22). Likewise, functional testing of these features in other miRNA targets may further shed light on the role factors beyond complementarity play in miRNA-mediated regulation.

3.4 Material and Methods

3.4.1 Bioinformatics workflow to identify HE and LE targets across species

Mature miRNA sequences for all species were retrieved from miRBase v22 (Kozomara et al., 2019). Where multiple isomiRs were found, the isomiR with the highest abundance found on a plant next-generation sequencing database (https://mpss.danforthcenter.org) was used (Nakano et al., 2020). The isomiR sequences analysed can be found in Table S7. Targets were predicted using psRNATarget v2 (Dai et al., 2018). Default settings were used for analysis except the expectation score which was decreased to 3 for all miRNAs except miR167, miR398 and miR408. An expectation score of 5 was used for these miRNAs as their genes from the VAT set exceeds an expectation score of 3. The resulting predicted targets were then analysed using Whole-Degradome-based Plant MicroRNA-Target Interaction Analysis Server (WPMIAS) (Fei et al., 2020). The "Advanced II" > "Use psRNATarget predicted results directly" option was used for analysis by WPMIAS for all miRNAs. Default settings were used for all miRNAs except for miR162, miR396, miR398 and miR408 where "Offset from spliced position (nt)" was set to 1 as the previously validated targets of these miRNAs can only be identified at this setting.

The transcriptome libraries from psRNATarget and WPMIAS used for analysis can be found on Table S8. Degradome data retrieved from WPMIAS was then used as input and analysed using TRUEE to identify HE and LE targets. Analysis by TRUEE was performed at a *Cleavage Tag Abundance* of ≥ 5 TP10M, *Library % Cut-off* of 20% and a *Target Category* of both Category 1 and 2 targets. R script used for this analysis is accessible on the Open Science Framework page for this project https://osf.io/3j65e/. Target Categories as defined in WPMIAS were used in this study (Fei et al., 2020).

3.4.2 Quantifying sequence conservation and RNA secondary structure prediction

For each target gene twenty sequences of homologues from diverse species of land plants were retrieved from nBLAST using the *A. thaliana* sequence as input. Diversity was achieved by choosing species ranging across major taxonomic divisions where homologues were available. Taxonomic divisions were eudicots-rosids, eudicots-asterids, eudicots-ranunculids, monocots, *Amborella trichopoda*, gynmnosperms, lycophytes and bryophytes. One sequence was chosen for each species with the highest sequence identity of the whole gene. Only up to two mismatches in the binding site were allowed for each sequence. Default settings were used for BLASTn.

Sequences were aligned using Multiple Alignment using Fast Fourier Transform (MAFFT) with default settings (Katoh & Standley, 2013). Conservation was determined using phyloP using LRT in “CON” conservation mode to measure slower than neutral evolution (Pollard et al., 2010) where a positive phyloP score denotes conservation. phyloP scores were generated using Phylogenetic Analysis with Space/Time Models (rPHAST) (Hubisz et al., 2011). As input into rPHAST, phylogenetic trees were generated using Simple Phylogeny to fit phylogenetic tree to the alignment and determine a neutral model (https://www.ebi.ac.uk/Tools/phylogeny/simple_phylogeny/) (Larkin et al., 2007; Goujon et al., 2010; McWilliam et al., 2013, Madeira et al., 2019). The neutral model is when the changes of the sequence is under neutral genetic drift. Hence by comparing the substitution rate at a particular nucleotide position to the neutral model, whether this nucleotide is conserved or undergoing accelerated substitution can be determined.

phyloP scores were further adjusted for FDR using the “BH” method (Benjamini & Hochberg, 1995). In this study, Individual nucleotide positions were only considered conserved if they possessed an FDR-adjusted phyloP score of ≥ 1.0. Sequences were considered conserved if conserved nucleotides occurred ≥ 4 in a row. The R script used to calculate phyloP score is accessible on the Open Science Framework page for this project https://osf.io/3j65e/.

The consensus RNA secondary structure was analysed using RNAfold (Hofacker et al., 1994; Turner et al., 2009) from a sequence window consisting of approximately 50 nts upstream and downstream of the binding site (100 nt + 21 nt + 100 nt = 221 nt window). Default parameters were used except temperature which was set at 22 ^{o}C to reflect Arabidopsis growth temperatures.

3.4.3 Identification of the presence of conserved sequence in HE and LE targets across species

Having identified conserved sequences flanking miRNA binding sites, the HE and LE targets of the primary target family were then analysed for the presence of these conserved sequences (Figure S3, Figure S4). Although well-known to be targeted by miRNAs, *TAS3* and *IPS1* was not included in subsequent analysis as no degradome data was available (Franco-Zorilla et al., 2007; Allen et al., 2005).

Transcript sequences used to identify the presence of the conserved sequences for HE and LE targets of miRNA-target modules [miR159:MYB33; miR160:*ARF10* and *ARF17*; miR171:*HAM*; miR319:*TCP2* and*TCP4*; miR396:*GRF3*; TasiARF:*ARF2*, *ARF3* and *ARF4*] were retrieved from transcriptomes downloaded from the Genome portal of the Department of Energy Joint Genome Institute (Grigoriev et al., 2012; Nordberg et al., 2014) (https://phytozome.jgi.doe.gov/pz/portal.html) (Goodstein et al., 2012) (Table S9). The presence of the conserved sequences was identified using an in-house R script which is accessible on the Open Science Framework page for this project https://osf.io/3j65e/. The workflow is described in Figure 3.25.

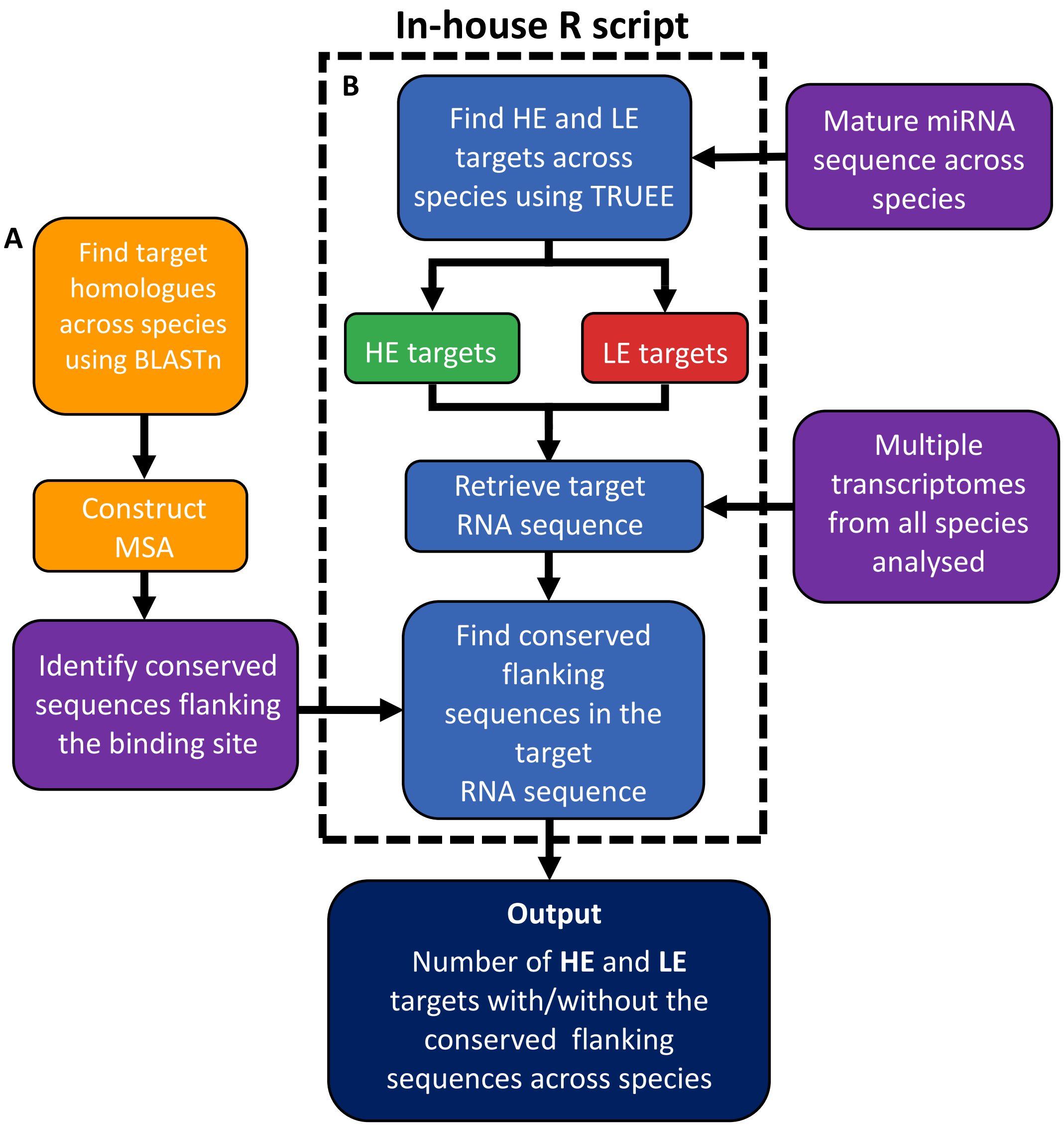


Figure 3.25. The workflow for identifying the number of HE and LE targets with/without the conserved flanking sequences A) For each miRNA-target module analysed, target homologues across species were found using BLASTn and used to construct an MSA to identify conserved sequences flanking the miRNA binding site. B) Mature miRNA sequences across species were used as input into TRUEE to identify HE and LE targets. The target RNA sequences was retrieved from transcriptome input from all species analysed. All retrieved RNA sequences were analysed for the conserved sequences identified in step A. The workflow output is the number of HE and LE targets with and without the conserved flanking sequences across all species analysed.

3.4.4 Data visualization

MSAs were visualised using Jalview (Waterhouse et al., 2009). The consensus sequence around the binding site, including the conserved sequences, was used to generate sequence logos using WebLogo (Crooks et al., 2004). The consensus RNA secondary structure was analysed using RNAalifold (Bernhart et al., 2008) from a sequence window consisting of approximately 50 nts upstream and downstream of the binding site (50 nt + 21 nt + 50 nt = 121 nt window). This window was extended to approximately 100 nts upstream and downstream of the binding site for miR171:*HAM* where the conservation of sequences flanking the binding site appeared to extend beyond a window of 121 nts. Default parameters were used except temperature which was set at 22 °C to generally reflect plant growth temperatures. T-plots of miRNA targets were adapted from WPMIAS (Fei et al., 2020). All graphs were generated using the R package, ggplot2, except for the pie charts which were generated using Excel.

3.4.5 PANTHER ID acquisition

PANTHER IDs, which were used to sort HE and LE targets into gene families, from Phytozome v12 via Phytomine, the InterMine interface to Phytozome (https://phytozome.jgi.doe.gov/phytomine/begin.do) (Goodstein et al., 2012; Smith et al., 2012; Mi et al., 2013). Phytomine was accessed using IntermineR, an R package providing an R interface with InterMine-Powered Databases, and was incorporated into the R-script (Kyritsis et al., 2019). Genes with no associated PANTHER ID are excluded from analysis and is the cause of discrepancy in the numbers between analysis using (Figure 3.2) and not using (Figure 3.1) PANTHER IDs. Analysis was performed using Phytozome v12 which has been made obsolete and replaced, and therefore results may differ.

3.4.6 Generation of *ARF10* entry clones using Gateway™ cloning (BP reaction)

The *ARF10* gene sequence was amplified from genomic DNA using primers with *attb1* and *attb2* sites for Gateway™ cloning (Invitrogen™) (Table S10). All procedures were performed as per manufacturer's protocol unless otherwise stated. PCR amplification was performed using high fidelity KOD Hot Start DNA Polymerase (Novagen™) using the following cycling conditions: 1 cycle of 95 °C for 2 min, 35 cycles of 95 °C/20 sec, 55 °C/10 sec, 70 °C for 15 sec/kb extension time according to amplicon size, and 1 cycle of 70 °C for 10 min. PCR products were analysed by agarose gel electrophoresis. Products corresponding to the expected amplicon sizes were excised from the gel and purified using the Wizard® SV Gel and PCR Clean-Up System (Promega).

PCR products were cloned into the donor vector, pDONOR/zeo (Invitrogen), using the Gateway™ BP Clonase™ II enzyme mix (Invitrogen™) to produce an *ARF10* (henceforth *ARF10-WT*) entry clone. The resulting reaction was transformed into Alpha-Select Gold Efficiency competent *E. coli* cells (Bioline) by heat shock and recovered in low-salt Luria Broth (LB) at 37 °C for 1 hr. *E. coli* were cultured on low-salt LB agar plates containing 50 µg/mL Zeocin™ (Invitrogen™) over night at 37 °C. Positive clones were sub-cultured overnight at 37 °C in LB with Zeocin™. Plasmids were extracted using the FavorPrep™ Plasmid Extraction Mini Kit (Favorgen®). Plasmids were screened using diagnostic restriction enzyme digestion. Sequences were then verified via Sanger sequencing with the M13 forward and reverse primers using the ABI PRISM® BigDye® Terminator v3.1 cycle sequencing kit (Applied Biosystems™). Sequenced products were then purified using Wizard® SV Gel and PCR Clean-Up System (Promega). Purified sequences were then precipitated and analysed at John Curtin School of Medical Research, Australian National University, Canberra.

3.4.7 Site-directed mutagenesis

The *ARF10-FM* and *rmARF10* entry clones were generated by introducing mutations to *ARF10-WT* entry clone using site-directed mutagenesis (Liu & Naismith, 2009). For both *ARF10* variants, primer pairs for site-directed mutagenesis were designed with complementary overlapping regions at the 3’ end where the mutations were situated (Table S10). Primers also contained non-overlapping sequences at the 5’ end to minimize primer dimerisation and allow primers to use the PCR product as a template. Non-overlapping sequences also possessed a 5-10 °C higher Tm and were longer than the overlapping regions to promote annealing to the plasmid template over primer dimerisation. PCR was performed using high fidelity KOD Hot Start DNA Polymerase (Novagen) using the following cycling conditions: 1 cycle of 95 °C for 2 min, 35 cycles of 95 °C /20 sec, 55 °C /10 sec, 70 °C for 100 sec/kb extension time according to amplicon size, and 1 cycle of 70 °C for 10 min. The PCR product was digested with 2 µL DpnI at 37 °C for 4 hr to remove the unmutated parental vector. The PCR reaction was purified using the Wizard® SV Gel and PCR Clean-Up System (Promega). The *ARF10-FM* and *rmARF10* entry clones were transformed into *DH5-α E. coli* using electroporation and recovered in low-salt LB at 37 °C for 1 hr. *E. coli* were cultured on LB agar plates containing 50 µg/mL Zeocin™ over night at 37 °C. Plasmids were extracted and confirmed using restriction enzyme digestion analysis and sequencing as above.

3.4.8 Generation of *ARF10* expression clones using Gateway™ cloning (LR reaction)

The correct entry clones were sub-cloned into the destination vector, pGWB602Ω (Nakamura et al., 2010), using the Gateway™ LR Clonase™ II enzyme mix (Invitrogen™) to generate the *ARF10-WT*, *ARF10-FM* and *rmARF10* expression clones. Expression clones were transformed into *DH5-α E. coli* using electroporation and recovered in LB at 37 °C for 1 hr. *E. coli* were cultured on LB agar plates containing 50 µg/mL Spectinomycin over night at 37 °C. Positive clones were sub-cultured overnight at 37 °C in LB with Spectinomycin. Plasmids were extracted and confirmed using restriction enzyme digestion analysis.

3.4.9 Transformation of Agrobacteria

Expression clones were transformed into a GV3101 strain of *Agrobacterium tumefaciens Agrobacterium tumefaciens* by electroporation (Hellens et al., 2000) and recovered in LB at 28 °C for 4 hr. *Agrobacterium tumefaciens* were cultured on LB agar plates containing 50 ug/mL Rifamycin, 25 ug/mL Gentamicin and 50 ug/mL Spectinomycin at 28 °C for 48 hr. Single colonies were picked and used to inoculate 15 mL LB with the same antibiotics and temperature for 18-20 hr. Plasmids were extracted and confirmed using restriction enzyme digestion analysis.

3.4.10 Plant Material and Growth Conditions

Arabidopsis thaliana ecotype Columbia-0 (Col-0) plants were used in all experiments. Seeds of primary transformants were harvested and vapour sterilized using chlorine gas for 3-4 hr in a desiccator jar. Chlorine gas was generated by mixing 100 mL of 100% sodium hypochlorite with 3 mL of 36 % hydrochloric acid. Seeds were grown on soil (Debco® plugger soil mix with 3.5g/L Osmocote® fertiliser) or on Murashige and Skoog (MS) medium agar plates. All seeds were stratified for 24 hr at 2 °C in the dark. Seeds were then grown at 22 °C, 150-200 umol/m^2/sec light intensity, under 10hr day/12hr night conditions.

3.4.11 Transformation of Arabidopsis

Agrobacterium tumefaciens transformed with the *ARF10* variant expression clones were each inoculated into 15 mL LB with the appropriate antibiotics (above) and incubated for 18-20 hr at 28 °C. 1 mL of the liquid culture was inoculated into 250 mL LB with 25 µg/mL Gentamicin and 50 µg/mL Spectinomycin and incubated at 28 °C for 48 hr with constant shaking at 220 rpm. To prepare the culture used to transform Arabidopsis, *Agrobacterium tumefaciens* cultures were centrifuged at 5000 rpm for 15 mL and resuspended in infiltration media containing 5% sucrose and 0.03% of the surfactant, Silwet L-77® (Clough and Bent, 1998). Arabidopsis was transformed by dipping the inflorescences into the infiltration culture for 30 sec. Plants were covered in plastic bags and kept in the dark for 24 hr before being returned to growth chambers. An *Agrobacterium tumefaciens* with the pGWB602Ω empty vector was also transformed into Arabidopsis and used as a transgenic control.

Primary transformant seeds were then harvested and sterilised as above. For the selection of transformants, seeds were sown on agar plates containing 0.5X MS agar plates with the appropriate selective antibiotic, BASTA. Seeds were stratified and grown as above. Primary transformants were identified at 6-7 days old and transplanted onto soil.

3.4.12 Statistical analysis

ANOVA was used to analyse the average Expectation Score between HE targets and LE targets. Plant morphological phenotyping results were analysed using Pearson's Chi-square test.

References

Abdel-Ghany, S. E., & Pilon, M. (2008). MicroRNA-mediated systemic down-regulation of copper protein expression in response to low copper availability in Arabidopsis. *Journal of Biological Chemistry*, 283(23), 15932–15945. https://doi.org/10.1074/jbc.M801406200

Allen, E., Xie, Z., Gustafson, A. M., & Carrington, J. C. (2005). microRNA-directed phasing during trans-acting siRNA biogenesis in plants. *Cell*, 121(2), 207–221. https://doi.org/10.1016/j.cell.2005.04.004

Amborella Genome Project (2013). The Amborella genome and the evolution of flowering plants. *Science* (New York, N.Y.), 342(6165), 1241089. https://doi.org/10.1126/science.1241089

Axtell, M. J., & Meyers, B. C. (2018). Revisiting Criteria for Plant MicroRNA Annotation in the Era of Big Data. *The Plant Cell*, 30(2), 272–284. https://doi.org/10.1105/tpc.17.00851

Banks, J. A., Nishiyama, T., Hasebe, M., Bowman, J. L., Gribskov, M., dePamphilis, C., Albert, V. A., Aono, N., Aoyama, T., Ambrose, B. A., Ashton, N. W., Axtell, M. J., Barker, E., Barker, M. S., Bennetzen, J. L., Bonawitz, N. D., Chapple, C., Cheng, C., Correa, L. G., Dacre, M., … Grigoriev, I. V. (2011). The Selaginella genome identifies genetic changes associated with the evolution of vascular plants. *Science* (New York, N.Y.), 332(6032), 960–963. https://doi.org/10.1126/science.1203810

Bari, A., Orazova, S., & Ivashchenko, A. (2013). miR156- and miR171-binding sites in the protein-coding sequences of several plant genes. *BioMed research international*, 2013, 307145. https://doi.org/10.1155/2013/307145

Bazin, J., Khan, G. A., Combier, J.-P., Bustos-Sanmamed, P., Debernardi, J. M., Rodriguez, R., Sorin, C., Palatnik, J., Hartmann, C., Crespi, M., & Lelandais-Brière, C. (2013). miR396 affects mycorrhization and root meristem activity in the legume Medicago truncatula. *The Plant Journal*, 74(6), 920–934. https://doi.org/https://doi.org/10.1111/tpj.12178

Beier, S., Himmelbach, A., Colmsee, C., Zhang, X. Q., Barrero, R. A., Zhang, Q., Li, L., Bayer, M., Bolser, D., Taudien, S., Groth, M., Felder, M., Hastie, A., Šimková, H., Staňková, H., Vrána, J., Chan, S., Muñoz-Amatriaín, M., Ounit, R., Wanamaker, S., … Mascher, M. (2017). Construction of a map-based reference genome sequence for barley, Hordeum vulgare L. *Scientific data*, 4, 170044. https://doi.org/10.1038/sdata.2017.44

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society*. Series B (Methodological), 57(1), 289–300. http://www.jstor.org/stable/2346101

Bernhart, S. H., Hofacker, I. L., Will, S., Gruber, A. R., & Stadler, P. F. (2008). RNAalifold: improved consensus structure prediction for RNA alignments. *BMC Bioinformatics*, 9(1), 474. https://doi.org/10.1186/1471-2105-9-474

Boccara, M., Sarazin, A., Thiébeauld, O., Jay, F., Voinnet, O., Navarro, L., & Colot, V. (2014). The Arabidopsis miR472-RDR6 Silencing Pathway Modulates PAMP- and Effector-Triggered Immunity through the Post-transcriptional Control of Disease Resistance Genes. *PLOS Pathogens*, 10(1), e1003883. https://doi.org/10.1371/journal.ppat.1003883

Bonnet, E., He, Y., Billiau, K., & Van de Peer, Y. (2010). TAPIR, a web server for the prediction of plant microRNA targets, including target mimics. *Bioinformatics*, 26(12), 1566–1568. https://doi.org/10.1093/bioinformatics/btq233

Brousse, C., Liu, Q., Beauclair, L., Deremetz, A., Axtell, M. J., & Bouché, N. (2014). A non-canonical plant microRNA target site. *Nucleic Acids Research*, 42(8), 5270–5279. https://doi.org/10.1093/nar/gku157

Chauhan, S., Yogindran, S., & Rajam, M. V. (2017). Role of miRNAs in biotic stress reactions in plants. *Indian Journal of Plant Physiology*, 22(4), 514–529. https://doi.org/10.1007/s40502-017-0347-3

Cheng, F., Pan, Y., Lu, Y.-M., Zhu, L., & Chen, S. (2017). RNA-Binding Protein Dnd1 Promotes Breast Cancer Apoptosis by Stabilizing the Bim mRNA in a miR-221 Binding Site. *BioMed Research International*, 2017, 9596152. https://doi.org/10.1155/2017/9596152

Chorostecki, U., Crosa, V. A., Lodeyro, A. F., Bologna, N. G., Martin, A. P., Carrillo, N., Schommer, C., & Palatnik, J. F. (2012). Identification of new microRNA-regulated genes by conserved targeting in plant species. *Nucleic Acids Research*, 40(18), 8893–8904. https://doi.org/10.1093/nar/gks625

Clough, S. J., & Bent, A. F. (1998). Floral dip: a simplified method for Agrobacterium -mediated transformation of Arabidopsis thaliana. *The Plant Journal*, 16(6), 735–743. https://doi.org/https://doi.org/10.1046/j.1365-313x.1998.00343.x

Crooks, G. E., Hon, G., Chandonia, J.-M., & Brenner, S. E. (2004). WebLogo: a sequence logo generator. *Genome Research*, 14(6), 1188–1190. https://doi.org/10.1101/gr.849004

Dai, X., & Zhao, P. X. (2011). PsRNATarget: A plant small RNA target analysis server. *Nucleic Acids Research*, 39, 155–159. https://doi.org/10.1093/nar/gkr319

Dai, X., Zhuang, Z., & Zhao, P. X. (2018). PsRNATarget: A plant small RNA target analysis server (2017 release). *Nucleic Acids Research*, 46(W1), W49–W54. https://doi.org/10.1093/nar/gky316

Debernardi, J. M., Rodriguez, R. E., Mecchia, M. A., & Palatnik, J. F. (2012). Functional specialization of the plant miR396 regulatory network through distinct microRNA-target interactions. *PLoS Genetics*, 8(1). https://doi.org/10.1371/journal.pgen.1002419

Deveson, I., Li, J., & Millar, A. A. (2013). MicroRNAs with analogous target complementarities perform with highly variable efficacies in Arabidopsis. *FEBS Letters*, 587(22), 3703–3708. https://doi.org/https://doi.org/10.1016/j.febslet.2013.09.037

Fei, Y., Mao, Y., Shen, C., Wang, R., Zhang, H., & Huang, J. (2020). WPMIAS: Whole-degradome-based Plant MicroRNA–target Interaction Analysis Server. *Bioinformatics*, 36(6), 1937–1939. https://doi.org/10.1093/bioinformatics/btz820

Franco-Zorrilla, J. M., Valli, A., Todesco, M., Mateos, I., Puga, M. I., Rubio-Somoza, I., Leyva, A., Weigel, D., García, J. A., & Paz-Ares, J. (2007). Target mimicry provides a new mechanism for regulation of microRNA activity. *Nature Genetics*, 39(8), 1033–1037. https://doi.org/10.1038/ng2079

Gawroński, P., Jensen, P. E., Karpiński, S., Leister, D., & Scharff, L. B. (2018). Pausing of Chloroplast Ribosomes Is Induced by Multiple Features and Is Linked to the Assembly of Photosynthetic Complexes. *Plant Physiology*, 176(3), 2557–2569. https://doi.org/10.1104/pp.17.01564

German, M. a, Pillay, M., Jeong, D.-H., Hetawal, A., Luo, S., Janardhanan, P., Kannan, V., Rymarquis, L. a, Nobuta, K., German, R., De Paoli, E., Lu, C., Schroth, G., Meyers, B. C., & Green, P. J. (2008). Global identification of microRNA-target RNA pairs by parallel analysis of RNA ends. *Nature Biotechnology*, 26(8), 941–946. https://doi.org/10.1038/nbt1417

Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., & Rokhsar, D. S. (2012). Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Research*, 40(D1), D1178–D1186. https://doi.org/10.1093/nar/gkr944

Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J., & Lopez, R. (2010). A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Research*, 38(Web Server issue), W695–W699. https://doi.org/10.1093/nar/gkq313

Grigoriev, I. V., Nordberg, H., Shabalov, I., Aerts, A., Cantor, M., Goodstein, D., Kuo, A., Minovitsky, S., Nikitin, R., Ohm, R. A., Otillar, R., Poliakov, A., Ratnere, I., Riley, R., Smirnova, T., Rokhsar, D., & Dubchak, I. (2012). The Genome Portal of the Department of Energy Joint Genome Institute. *Nucleic Acids Research*, 40(D1), 26–32. https://doi.org/10.1093/nar/gkr947

Gu, W., Wang, X., Zhai, C., Xie, X., & Zhou, T. (2012). Selection on synonymous sites for increased accessibility around mirna binding sites in plants. *Molecular Biology and Evolution*, 29(10), 3037–3044. https://doi.org/10.1093/molbev/mss109

Hellens, R. P., Allan, A. C., Friel, E. N., Bolitho, K., Grafton, K., Templeton, M. D., Karunairetnam, S., Gleave, A. P., & Laing, W. A. (2005). Transient expression vectors for functional genomics, quantification of promoter activity and RNA silencing in plants. *Plant Methods*, 1(1), 13. https://doi.org/10.1186/1746-4811-1-13

Hofacker, I. L., Fontana, W., Stadler, P. F., Bonhoeffer, L. S., Tacker, M., & Schuster, P. (1994). Fast folding and comparison of RNA secondary structures. *Monatshefte Für Chemie / Chemical Monthly*, 125(2), 167–188. https://doi.org/10.1007/BF00818163

Hubisz, M. J., Pollard, K. S., Siepel, A. (2011). PHAST and RPHAST: phylogenetic analysis with space/time models. *Briefings in Bioinformatics*, 12(1), 41-51. doi:10.1093/bib/bbq072

Humphries, B., & Yang, C. (2015). The microRNA-200 family: small molecules with novel roles in cancer development, progression and therapy. *Oncotarget*; Vol 6, No 9. https://www.oncotarget.com/article/3052/text/

International Brachypodium Initiative (2010). Genome sequencing and analysis of the model grass Brachypodium distachyon. *Nature*, 463(7282), 763–768. https://doi.org/10.1038/nature08747

International Peach Genome Initiative, Verde, I., Abbott, A. G., Scalabrin, S., Jung, S., Shu, S., Marroni, F., Zhebentyayeva, T., Dettori, M. T., Grimwood, J., Cattonaro, F., Zuccolo, A., Rossini, L., Jenkins, J., Vendramin, E., Meisel, L. A., Decroocq, V., Sosinski, B., Prochnik, S., Mitros, T., … Rokhsar, D. S. (2013). The high-quality draft genome of peach (Prunus persica) identifies unique patterns of genetic diversity, domestication and genome evolution. *Nature genetics*, 45(5), 487–494. https://doi.org/10.1038/ng.2586

Iwai, N., Yasui, K., Tomie, A., Gen, Y., Terasaki, K., Kitaichi, T., Soda, T., Yamada, N., Dohi, O., Seko, Y., Umemura, A., Nishikawa, T., Yamaguchi, K., Moriguchi, M., Konishi, H., Naito, Y., & Itoh, Y. (2018). Oncogenic miR-96-5p inhibits apoptosis by targeting the caspase-9 gene in hepatocellular carcinoma. *Int J Oncol*, 53(1), 237–245. https://doi.org/10.3892/ijo.2018.4369

Jaillon, O., Aury, J. M., Noel, B., Policriti, A., Clepet, C., Casagrande, A., Choisne, N., Aubourg, S., Vitulo, N., Jubin, C., Vezzi, A., Legeai, F., Hugueney, P., Dasilva, C., Horner, D., Mica, E., Jublot, D., Poulain, J., Bruyère, C., Billault, A., … French-Italian Public Consortium for Grapevine Genome Characterization (2007). The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature*, 449(7161), 463–467. https://doi.org/10.1038/nature06148

Jones-Rhoades, M. W., & Bartel, D. P. (2004). Computational Identification of Plant MicroRNAs and Their Targets, Including a Stress-Induced miRNA. *Molecular Cell*, 14(6), 787–799. https://doi.org/10.1016/j.molcel.2004.05.027

Jones-Rhoades, M. W., Bartel, D. P., & Bartel, B. (2006). MicroRNAs and Their Regulatory Roles in Plants. *Plant Biology*, 1–16. https://doi.org/10.1146/annurev.arplant.57.032905.105218

Jones-Rhoades, M. W. (2012). Conservation and divergence in plant microRNAs. *Plant Molecular Biology*, 80(1), 3–16. https://doi.org/10.1007/s11103-011-9829-2

Katoh, K., & Standley, D. M. (2013). MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution*, 30(4), 772–780. https://doi.org/10.1093/molbev/mst010

Kawashima, C. G., Yoshimoto, N., Maruyama-Nakashita, A., Tsuchiya, Y. N., Saito, K., Takahashi, H., & Dalmay, T. (2009). Sulphur starvation induces the expression of microRNA-395 and one of its target genes but in different cell types. *The Plant Journal*, 57(2), 313–321. https://doi.org/10.1111/j.1365-313X.2008.03690.x

Kedde, M., Strasser, M. J., Boldajipour, B., Vrielink, J. A. F. O., Slanchev, K., le Sage, C., Nagel, R., Voorhoeve, P. M., van Duijse, J., Ørom, U. A., Lund, A. H., Perrakis, A., Raz, E., & Agami, R. (2007). RNA-Binding Protein Dnd1 Inhibits MicroRNA Access to Target mRNA. *Cell*, 131(7), 1273–1286. https://doi.org/https://doi.org/10.1016/j.cell.2007.11.034

Kedde, M., van Kouwenhove, M., Zwart, W., Oude Vrielink, J. A. F., Elkon, R., & Agami, R. (2010). A Pumilio-induced RNA structure switch in p27-3’ UTR controls miR-221 and miR-222 accessibility. *Nature Cell Biology*, 12(10), 1014–1020. https://doi.org/10.1038/ncb2105

Kozomara, A., Birgaoanu, M., & Griffiths-Jones, S. (2019). miRBase: from microRNA sequences to function. *Nucleic Acids Research*, 47(D1), D155–D162. https://doi.org/10.1093/nar/gky1141

Kyritsis, K. A., Wang, B., Sullivan, J., Lyne, R., & Micklem, G. (2019). InterMineR: an R package for InterMine databases. *Bioinformatics*, 35(17), 3206–3207. https://doi.org/10.1093/bioinformatics/btz039

Lamesch, P., Berardini, T. Z., Li, D., Swarbreck, D., Wilks, C., Sasidharan, R., Muller, R., Dreher, K., Alexander, D. L., Garcia-Hernandez, M., Karthikeyan, A. S., Lee, C. H., Nelson, W. D., Ploetz, L., Singh, S., Wensel, A., & Huala, E. (2012). The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. *Nucleic acids research*, 40, D1202–D1210. https://doi.org/10.1093/nar/gkr1090

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., & Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics*, 23(21), 2947–2948. https://doi.org/10.1093/bioinformatics/btm404

Li, F., Zheng, Q., Vandivier, L. E., Willmann, M. R., Chen, Y., & Gregory, B. D. (2012). Regulatory impact of RNA secondary structure across the Arabidopsis transcriptome. *The Plant Cell*, 24(11), 4346–4359. https://doi.org/10.1105/tpc.112.104232

Li, J.-F., Chung, H. S., Niu, Y., Bush, J., McCormack, M., & Sheen, J. (2013). Comprehensive protein-based artificial microRNA screens for effective gene silencing in plants. *The Plant Cell*, 25(5), 1507–1522. https://doi.org/10.1105/tpc.113.112235

Li, J., Reichel, M., Li, Y., & Millar, A. A. (2014). The functional scope of plant microRNA-mediated silencing. *Trends in Plant Science*, 19(12), 750–756. https://doi.org/10.1016/j.tplants.2014.08.006

Liang, G., Yang, F., & Yu, D. (2010). MicroRNA395 mediates regulation of sulfate accumulation and allocation in Arabidopsis thaliana. *The Plant Journal*, 62(6), 1046–1057. https://doi.org/10.1111/j.1365-313X.2010.04216.x

Liang, Y., Tan, Z.-M., Zhu, L., Niu, Q.-K., Zhou, J.-J., Li, M., Chen, L.-Q., Zhang, X.-Q., & Ye, D. (2013). MYB97, MYB101 and MYB120 Function as Male Factors That Control Pollen Tube-Synergid Interaction in Arabidopsis thaliana Fertilization. *PLOS Genetics*, 9(11), e1003933. https://doi.org/10.1371/journal.pgen.1003933

Lin, W.-Y., Lin, Y.-Y., Chiang, S.-F., Syu, C., Hsieh, L.-C., & Chiou, T.-J. (2018). Evolution of microRNA827 targeting in the plant kingdom. *New Phytologist*, 217(4), 1712–1725. https://doi.org/https://doi.org/10.1111/nph.14938

Liu, P. P., Montgomery, T. A., Fahlgren, N., Kasschau, K. D., Nonogaki, H., & Carrington, J. C. (2007). Repression of AUXIN RESPONSE FACTOR10 by microRNA160 is critical for seed germination and post-germination stages. *The Plant Journal*, 52(1), 133–146. https://doi.org/10.1111/j.1365-313X.2007.03218.x

Liu, H., & Naismith, J. H. (2008). An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC Biotechnology*, 8, 91. https://doi.org/10.1186/1472-6750-8-91

Liu, Q., Wang, F., & Axtell, M. J. (2014). Analysis of complementarity requirements for plant MicroRNA targeting using a Nicotiana benthamiana quantitative transient assay. *The Plant Cell*, 26(2), 741–753. https://doi.org/10.1105/tpc.113.120972

Lustig, Y., Barhod, E., Ashwal-Fluss, R., Gordin, R., Shomron, N., Baruch-Umansky, K., Hemi, R., Karasik, A., & Kanety, H. (2014). RNA-Binding Protein PTB and MicroRNA-221 Coregulate AdipoR1 Translation and Adiponectin Signaling. *Diabetes*, 63(2), 433–445. https://doi.org/10.2337/db13-1032

Madeira, F., Park, Y. mi, Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A. R. N., Potter, S. C., Finn, R. D., & Lopez, R. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Research*, 47(W1), W636–W641. https://doi.org/10.1093/nar/gkz268

Magenta, A., Ciarapica, R., & Capogrossi, M. C. (2017). The Emerging Role of miR-200 Family in Cardiovascular Diseases. *Circulation Research*, 120(9), 1399–1402. https://doi.org/10.1161/CIRCRESAHA.116.310274

Mallory, A. C., Reinhart, B. J., Jones-Rhoades, M. W., Tang, G., Zamore, P. D., Barton, M. K., & Bartel, D. P. (2004). MicroRNA control of PHABULOSA in leaf development: importance of pairing to the microRNA 5’ region. *The EMBO Journal*, 23(16), 3356–3364. https://doi.org/10.1038/sj.emboj.7600340

Mallory, A. C., Bartel, D. P., & Bartel, B. (2005). MicroRNA-Directed Regulation of Arabidopsis AUXIN RESPONSE FACTOR17 Is Essential for Proper Development and Modulates Expression of Early Auxin Response Genes. *Development*, 17(May), 1–16. https://doi.org/10.1105/tpc.105.031716.1

McWilliam, H., Li, W., Uludag, M., Squizzato, S., Park, Y. M., Buso, N., Cowley, A. P., & Lopez, R. (2013). Analysis Tool Web Services from the EMBL-EBI. *Nucleic Acids Research*, 41(Web Server issue), W597–W600. https://doi.org/10.1093/nar/gkt376

Mi, H., Muruganujan, A., Casagrande, J. T., & Thomas, P. D. (2013). Large-scale gene function analysis with the PANTHER classification system. *Nature Protocols*, 8(8), 1551–1566. https://doi.org/10.1038/nprot.2013.092

Millar, A. A., & Gubler, F. (2005). The Arabidopsis GAMYB-Like Genes, MYB33 and MYB65, Are MicroRNA-Regulated Genes That Redundantly Facilitate Anther Development. *The Plant Cell*, 17(3), 705–721. https://doi.org/10.1105/tpc.104.027920

Morel, J.-B., Godon, C., Mourrain, P., Béclin, C., Boutet, S., Feuerbach, F., Proux, F., & Vaucheret, H. (2002). Fertile Hypomorphic ARGONAUTE (ago1) Mutants Impaired in Post-Transcriptional Gene Silencing and Virus Resistance. *The Plant Cell*, 14(3), 629–639. https://doi.org/10.1105/tpc.010358

Nakamura, S., Mano, S., Tanaka, Y., Ohnishi, M., Nakamori, C., Araki, M., Niwa, T., Nishimura, M., Kaminaka, H., Nakagawa, T., Sato, Y., & Ishiguro, S. (2010). Gateway binary vectors with the bialaphos resistance gene, bar, as a selection marker for plant transformation. Bioscience, biotechnology, and biochemistry, 74(6), 1315–1319. https://doi.org/10.1271/bbb.100184

Nakano, M., McCormick, K., Demirci, C., Demirci, F., Gurazada, S. G. R., Ramachandruni, D., Dusia, A., Rothhaupt, J. A., & Meyers, B. C. (2020). Next-Generation Sequence Databases: RNA and Genomic Informatics Resources for Plants. *Plant Physiology*, 182(1), 136–146. https://doi.org/10.1104/pp.19.00957

Naya, L., Paul, S., Valdés-López, O., Mendoza-Soto, A. B., Nova-Franco, B., Sosa-Valencia, G., Reyes, J. L., & Hernández, G. (2014). Regulation of Copper Homeostasis and Biotic Interactions by MicroRNA 398b in Common Bean. *PLOS ONE*, 9(1), e84416. https://doi.org/10.1371/journal.pone.0084416

Nordberg, H., Cantor, M., Dusheyko, S., Hua, S., Poliakov, A., Shabalov, I., Smirnova, T., Grigoriev, I. V., & Dubchak, I. (2014). The genome portal of the Department of Energy Joint Genome Institute: 2014 updates. Nucleic acids research, 42(Database issue), D26–D31. https://doi.org/10.1093/nar/gkt1069

Ouyang, S., Zhu, W., Hamilton, J., Lin, H., Campbell, M., Childs, K., Thibaud-Nissen, F., Malek, R. L., Lee, Y., Zheng, L., Orvis, J., Haas, B., Wortman, J., & Buell, C. R. (2007). The TIGR Rice Genome Annotation Resource: improvements and new features. *Nucleic acids research*, 35(Database issue), D883–D887. https://doi.org/10.1093/nar/gkl976

Palatnik, J. F., Wollmann, H., Schommer, C., Schwab, R., Boisbouvier, J., Rodriguez, R., Warthmann, N., Allen, E., Dezulian, T., Huson, D., Carrington, J. C., & Weigel, D. (2007). Sequence and Expression Differences Underlie Functional Specialization of Arabidopsis MicroRNAs miR159 and miR319. *Developmental Cell*, 13(1), 115–125. https://doi.org/https://doi.org/10.1016/j.devcel.2007.04.012

Park, Y. J., Lee, H. J., Kwak, K. J., Lee, K., Hong, S. W., & Kang, H. (2014). MicroRNA400-Guided Cleavage of Pentatricopeptide Repeat Protein mRNAs Renders Arabidopsis thaliana More Susceptible to Pathogenic Bacteria and Fungi. *Plant and Cell Physiology*, 55(9), 1660–1668. https://doi.org/10.1093/pcp/pcu096

Pilon, M. (2017). The copper microRNAs. *New Phytologist*, 213(3), 1030–1035. https://doi.org/https://doi.org/10.1111/nph.14244

Pollard, K. S., Hubisz, M. J., Rosenbloom, K. R., & Siepel, A. (2010). Detection of nonneutral substitution rates on mammalian phylogenies. *Genome research*, 20(1), 110–121. doi:10.1101/gr.097857.109

Rhoades, M. W., Reinhart, B. J., Lim, L. P., Burge, C. B., Bartel, B., & Bartel, D. P. (2002). Prediction of plant microRNA targets. *Cell*, 110(4), 513–520. https://doi.org/10.1016/S0092-8674(02)00863-2

Samad, A. F. A., Sajad, M., Nazaruddin, N., Fauzi, I. A., Murad, A. M. A., Zainal, Z., & Ismail, I. (2017). MicroRNA and Transcription Factor: Key Players in Plant Regulatory Network. *In Frontiers in Plant Science* (Vol. 8). https://www.frontiersin.org/article/10.3389/fpls.2017.00565

Schiefthaler, U., Balasubramanian, S., Sieber, P., Chevalier, D., Wisman, E., & Schneitz, K. (1999). Molecular analysis of NOZZLE, a gene involved in pattern formation and early sporogenesis during sex organ development in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences of the United States of America*, 96(20), 11664–11669. https://doi.org/10.1073/pnas.96.20.11664

Schmutz, J., Cannon, S. B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., Hyten, D. L., Song, Q., Thelen, J. J., Cheng, J., Xu, D., Hellsten, U., May, G. D., Yu, Y., Sakurai, T., Umezawa, T., Bhattacharyya, M. K., Sandhu, D., Valliyodan, B., Lindquist, E., … Jackson, S. A. (2010). Genome sequence of the palaeopolyploid soybean. *Nature*, 463(7278), 178–183. https://doi.org/10.1038/nature08670

Schnable, P. S., Ware, D., Fulton, R. S., Stein, J. C., Wei, F., Pasternak, S., Liang, C., Zhang, J., Fulton, L., Graves, T. A., Minx, P., Reily, A. D., Courtney, L., Kruchowski, S. S., Tomlinson, C., Strong, C., Delehaunty, K., Fronick, C., Courtney, B., Rock, S. M., … Wilson, R. K. (2009).

The B73 maize genome: complexity, diversity, and dynamics. *Science* (New York, N.Y.), 326(5956), 1112–1115. https://doi.org/10.1126/science.1178534

Schwab, R., Palatnik, J. F., Riester, M., Schommer, C., Schmid, M., & Weigel, D. (2005). Specific effects of microRNAs on the plant transcriptome. *Developmental Cell*, 8(4), 517–527. https://doi.org/10.1016/j.devcel.2005.01.018

Sharma, D., Tiwari, M., Pandey, A., Bhatia, C., Sharma, A., & Trivedi, P. K. (2016). MicroRNA858 Is a Potential Regulator of Phenylpropanoid Pathway and Plant Development. *Plant Physiology*, 171(2), 944–959. https://doi.org/10.1104/pp.15.01831

Smith, R. N., Aleksic, J., Butano, D., Carr, A., Contrino, S., Hu, F., Lyne, M., Lyne, R., Kalderimis, A., Rutherford, K., Stepan, R., Sullivan, J., Wakeling, M., Watkins, X., & Micklem, G. (2012). InterMine: a flexible data warehouse system for the integration and analysis of heterogeneous biological data. *Bioinformatics*, 28(23), 3163–3165. https://doi.org/10.1093/bioinformatics/bts577

Sun, Y.-H., Lu, S., Shi, R., & Chiang, V. L. (2011). Computational Prediction of Plant miRNA Targets. RNAi and Plant Gene Function Analysis: Methods and Protocols (H. Kodama & A. Komamine (Eds.); pp. 175–186). *Humana Press*. https://doi.org/10.1007/978-1-61779-123-9_12

Sunkar, R., Kapoor, A., & Zhu, J.-K. (2006). Posttranscriptional Induction of Two Cu/Zn Superoxide Dismutase Genes in Arabidopsis Is Mediated by Downregulation of miR398 and Important for Oxidative Stress Tolerance. *The Plant Cell*, 18(8), 2051–2065. https://doi.org/10.1105/tpc.106.041673

Sunkar, R., Li, Y.-F., & Jagadeeswaran, G. (2012). Functions of microRNAs in plant stress responses. *Trends in Plant Science*, 17(4), 196–203. https://doi.org/https://doi.org/10.1016/j.tplants.2012.01.010

Tang, H., Krishnakumar, V., Bidwell, S., Rosen, B., Chan, A., Zhou, S., Gentzbittel, L., Childs, K. L., Yandell, M., Gundlach, H., Mayer, K. F., Schwartz, D. C., & Town, C. D. (2014). An improved genome release (version Mt4.0) for the model legume Medicago truncatula. *BMC genomics*, 15, 312. https://doi.org/10.1186/1471-2164-15-312

Tang, J., & Chu, C. (2017). MicroRNAs in crop improvement: fine-tuners for complex traits. *Nature Plants*, 3(7), 17077. https://doi.org/10.1038/nplants.2017.77

Tian, L., Cai, D., Zhuang, D., Wang, W., Wang, X., Bian, X., Xu, R., & Wu, G. (2020). miR-96-5p Regulates Proliferation, Migration, and Apoptosis of Vascular Smooth Muscle Cell Induced by Angiotensin II via Targeting NFAT5. *Journal of Vascular Research*, 57(2), 86–96. https://doi.org/10.1159/000505457

Tomato Genome Consortium (2012). The tomato genome sequence provides insights into fleshy fruit evolution. Nature, 485(7400), 635–641. https://doi.org/10.1038/nature11119

Turner, D. H., & Mathews, D. H. (2010). NNDB: the nearest neighbor parameter database for predicting stability of nucleic acid secondary structure. *Nucleic Acids Research*, 38(Database issue), D280–D282. https://doi.org/10.1093/nar/gkp892

Vahabi, M., Pulito, C., Sacconi, A., Donzelli, S., D'Andrea, M., Manciocco, V., Pellini, R., Paci, P., Sanguineti, G., Strigari, L., Spriano, G., Muti, P., Pandolfi, P. P., Strano, S., Safarian, S., Ganci, F., & Blandino, G. (2019). miR-96-5p targets PTEN expression affecting radio-chemosensitivity of HNSCC cells. *Journal of Experimental & Clinical Cancer Research*: CR, 38(1), 141. https://doi.org/10.1186/s13046-019-1119-x

Velasco, R., Zharkikh, A., Affourtit, J., Dhingra, A., Cestaro, A., Kalyanaraman, A., Fontana, P., Bhatnagar, S. K., Troggio, M., Pruss, D., Salvi, S., Pindo, M., Baldi, P., Castelletti, S., Cavaiuolo, M., Coppola, G., Costa, F., Cova, V., Dal Ri, A., Goremykin, V., … Viola, R. (2010). The genome of the domesticated apple (Malus × domestica Borkh.). *Nature genetics*, 42(10), 833–839. https://doi.org/10.1038/ng.654

Wang, J.-W., Czech, B., & Weigel, D. (2009). miR156-Regulated SPL Transcription Factors Define an Endogenous Flowering Pathway in Arabidopsis thaliana. *Cell*, 138(4), 738–749. https://doi.org/https://doi.org/10.1016/j.cell.2009.06.014

Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M., & Barton, G. J. (2009). Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, 25(9), 1189–1191. https://doi.org/10.1093/bioinformatics/btp033

Wu, G. A., Prochnik, S., Jenkins, J., Salse, J., Hellsten, U., Murat, F., Perrier, X., Ruiz, M., Scalabrin, S., Terol, J., Takita, M. A., Labadie, K., Poulain, J., Couloux, A., Jabbari, K., Cattonaro, F., Del Fabbro, C., Pinosio, S., Zuccolo, A., Chapman, J., … Rokhsar, D. (2014). Sequencing of diverse mandarin, pummelo and orange genomes reveals complex history of admixture during citrus domestication. *Nature biotechnology*, 32(7), 656–662. https://doi.org/10.1038/nbt.2906

Yamasaki, H., Abdel-Ghany, S. E., Cohu, C. M., Kobayashi, Y., Shikanai, T., & Pilon, M. (2007). Regulation of Copper Homeostasis by Micro-RNA in Arabidopsis*. *Journal of Biological Chemistry*, 282(22), 16369–16378. https://doi.org/https://doi.org/10.1074/jbc.M700138200

Yang, M., Woolfenden, H. C., Zhang, Y., Fang, X., Liu, Q., Vigh, M. L., Cheema, J., Yang, X., Norris, M., Yu, S., Carbonell, A., Brodersen, P., Wang, J., & Ding, Y. (2020). Intact RNA structurome reveals mRNA structure-mediated regulation of miRNA cleavage in vivo. *Nucleic Acids Research*, 48(15), 8767–8781. https://doi.org/10.1093/nar/gkaa577

Zhang, K., Shi, X., Zhao, X., Ding, D., Tang, J., & Niu, J. (2015). Investigation of miR396 and growth-regulating factor regulatory network in maize grain filling. *Acta Physiologiae Plantarum*, 37(2), 28. https://doi.org/10.1007/s11738-014-1767-6

Zheng, Z., Reichel, M., Deveson, I., Wong, G., Li, J., & Millar, A. A. (2017). Target RNA Secondary Structure Is a Major Determinant of miR159 Efficacy. *Plant Physiology*, 174(3), 1764–1778. https://doi.org/10.1104/pp.16.01898

Zhou, H.-Y., Wu, C.-Q., & Bi, E.-X. (2019). MiR-96-5p inhibition induces cell apoptosis in gastric adenocarcinoma. *World Journal of Gastroenterology*, 25(47), 6823–6834. https://doi.org/10.3748/wjg.v25.i47.6823

Chapter 4

General Discussion

Abbreviations

amiRNAs – artificial miRNAs

AP2 – APETELA2-LIKE

APS – ATP-SULFURYLASE

ARF – AUXIN RESPONSE FACTOR

Cat Score – category score

DMS – dimethyl sulphate sequencing

GRF – GROWTH REGULATORY FACTOR

HD-ZIPIII – CLASS III HOMEODOMAIN LEUCINE ZIPPER

HE – high evidence

LE – low evidence

MTIs – miRNA-target interactions

NF-YA – NUCLEAR TRANSCRIPTION FACTOR Y SUBUNIT ALPHA

siRNA – small interfering RNA

T-plots – Target-plots

TCP – TEOSINTE BRANCHED1, CYCLOIDEA, and PROLIFERATING CELL NUCLEAR ANTIGEN BINDING FACTOR

TRUEE – Targets Ranked Using Experimental Evidence

WPMIAS – Whole-Degradome-based Plant MicroRNA-Target Interaction Analysis Server

4.1 TRUEE provides a new scoring schema independent of miRNA-target binding site complementarity

A long-standing limitation of plant miRNA biology is the identification of functionally relevant miRNA targets. Many miRNA-target prediction tools result in long lists of 100s to 1000s of targets many for which there is no or little experimental evidence supporting the presence of miRNA-mediated regulation, suggesting the majority of predicted targets are likely false positives (Addo-Quaye et al., 2009; Folkes et al., 2012; Dai et al., 2018; Fei et al., 2020). Furthermore, the scoring schema of the most widely used prediction tools are developed based on binding site complementarity where miRNA targets are ranked by a mismatch score (Addo-Quaye et al., 2009; Dai et al., 2018). This assumes that higher complementarity equates to a greater chance of miRNA-mediated cleavage. However, there are many instances that appear inconsistent with this assumption and therefore ranking targets on complementarity can be misleading (Brousse et al., 2014; Liu et al., 2014; Zheng et al., 2017). In Chapter 2, Targets Ranked Using Experimental Evidence (TRUEE) was developed to filter and rank targets from the degradome-based miRNA target prediction tool, WPMIAS, into high evidence (HE) and low evidence (LE) of miRNA-mediated regulation. By applying stringent parameters, TRUEE filters for candidate genes with the most robust evidence as targets. The problem of ranking targets by complementarity is also circumvented in TRUEE as the category score (*Cat Score*) scoring schema is derived solely from degradome data based on the strength (*Target Category* and *Cleavage Tag Abundance*) and frequency of a target's target-plot (T-plots) across multiple degradome libraries. Although TRUEE uses psRNATarget or Whole-Degradome-based Plant MicroRNA-Target Interaction Analysis Server (WPMIAS) predicted targets as input, which have scoring schemas based on complementarity, the Expectation Score ultimately does not contribute to the *Cat Score*. This scoring schema is made more effective due to the large number of degradome libraries available on WPMIAS (Fei et al., 2020). WPMIAS can currently simultaneously analyse 61 Arabidopsis libraries; a number no other target prediction tool has reported to date. Furthermore, with the growing number of publicly available degradome experiments, the effectiveness of the *Cat Score* grows with the increases to frequency. As there is a minimum number of libraries a target's T-plot must occur in, TRUEE filters for miRNA-target interactions (MTIs) with the highest confidence. This is in contrast to the original study using WPMIAS by Fei et al. (2020) where the presence of a gene as a Category 1 or 2 target in only two Arabidopsis libraries was required to be considered a target. The robustness of TRUEE was demonstrated in that the canonical targets corresponded to high ranking targets. Therefore, this clearly shows *Cat Score* to be correlated with literature regarding the extent of miRNA-mediated regulation and that TRUEE is able to filter out and identify strong MTIs that have clear functional roles (Supplementary Table 4 & 5).

4.2 TRUEE supports a narrow functional scope of miRNA-mediated regulation in plants

As a proof of concept, TRUEE was applied to Arabidopsis to identify the scope of miRNA-mediated regulation, designated the Arabidopsis targetome. The resulting Arabidopsis targetome gives an estimation of the number of functional MTIs in a plant for the first time.

At present, two opposing views exists on the functional scope of miRNA-mediated regulation in plants. Some studies have proposed there to be potentially thousands of MTIs in plants with diverse functions (Lindow & Krogh, 2005; Lindow et al., 2007; Meng et al., 2011; Bülow et al., 2012; Fei et al., 2020). This notion stems from studies reporting that most miRNAs in a plant are lineage-specific (Fahlgren et al., 2010; Chávez Montes et al., 2014; Cui et al., 2017), that MTIs appear to be evolutionarily fluid (Smith et al., 2015) and from evidence that passenger miRNAs are functional (Reviewed in Liu et al., 2017). This was also aided by advances to sequencing technology over the last decade leading to the identification and annotation of a multitude of low abundance young miRNAs which are then uploaded onto miRBase (Kozomara et al., 2019). Furthermore, supporting this is the number of targets predicted from current bioinformatic predictions which suggests the possibility that there be a multitude of miRNAs regulating a multitude of diverse target genes (Bülow et al., 2012; Dai et al., 2018; Fei et al., 2020).

Alternatively, others have proposed a much narrower functional scope of miRNA-mediated regulation (Meng et al., 2012; Li et al., 2014; Taylor et al., 2014; Taylor et al., 2017; Axtell & Meyers, 2018). Many publications have questioned the quality and validity of these miRNA entries on miRBase which are mostly user-submitted, and have suggested the greater majority of entries are false positives (Meng et al., 2012; Taylor et al., 2014; Taylor et al., 2017; Axtell & Meyers, 2018; Kozomara et al., 2019). Rather, it is proposed that a majority of non-conserved miRNAs are evolutionarily transient having no targets and play no functional role in the plant (Axtell, 2008; Cuperus et al., 2011).

Supporting the narrower view of functional MTIs, TRUEE failed to identify HE targets for the vast majority of the Arabidopsis miRNA entries in miRBase (Figure 4.1; Table 2.3). This lends support to the notion that young potential miRNAs frequently emerge which provides a large pool from which new MTIs of functional significance can be acquired (Rajagopalan et al., 2006; Fahlgren et al., 2007; Axtell et al., 2007; Axtell, 2008). However, this is rare and most young miRNAs remain targetless and non-functional and are undergoing neutral drift. Instead, the Arabidopsis miRNome largely consisted of a relatively small subset of conserved guide miRNAs (~25) which regulate the majority of targets in the Arabidopsis targetome (Figure 2.5). Furthermore, the strength of a MTI also correlated with its conservation in that the highest *Cat scores* corresponded with the most conserved canonical targets whereas very few less-conserved and uncharacterised targets had high *Cat scores* (Table S3-S6). Even within the conserved guide miRNAs, the conserved targets had the highest *Cat scores* (Figure 2.6; Table S4). As such, the targetome appears to predominantly consist of the previously characterised highly conserved MTIs.

These results were further supported in Chapter 3 where TRUEE was used to analyse 20 highly conserved miRNAs and the highly conserved tasiARF from degradomes across diverse plant species. For most miRNAs and the tasiARF, targets were mainly or even nearly exclusively homologues from one conserved gene family which corresponded to those found in the Arabidopsis targetome (Figure 3.2). Comparatively few additional targets were identified. Together, these results support the notion that most of the functional MTIs have already been identified and contradicts the idea that there may be 100s and 1000s of functional MTIs (Figure 2.5).

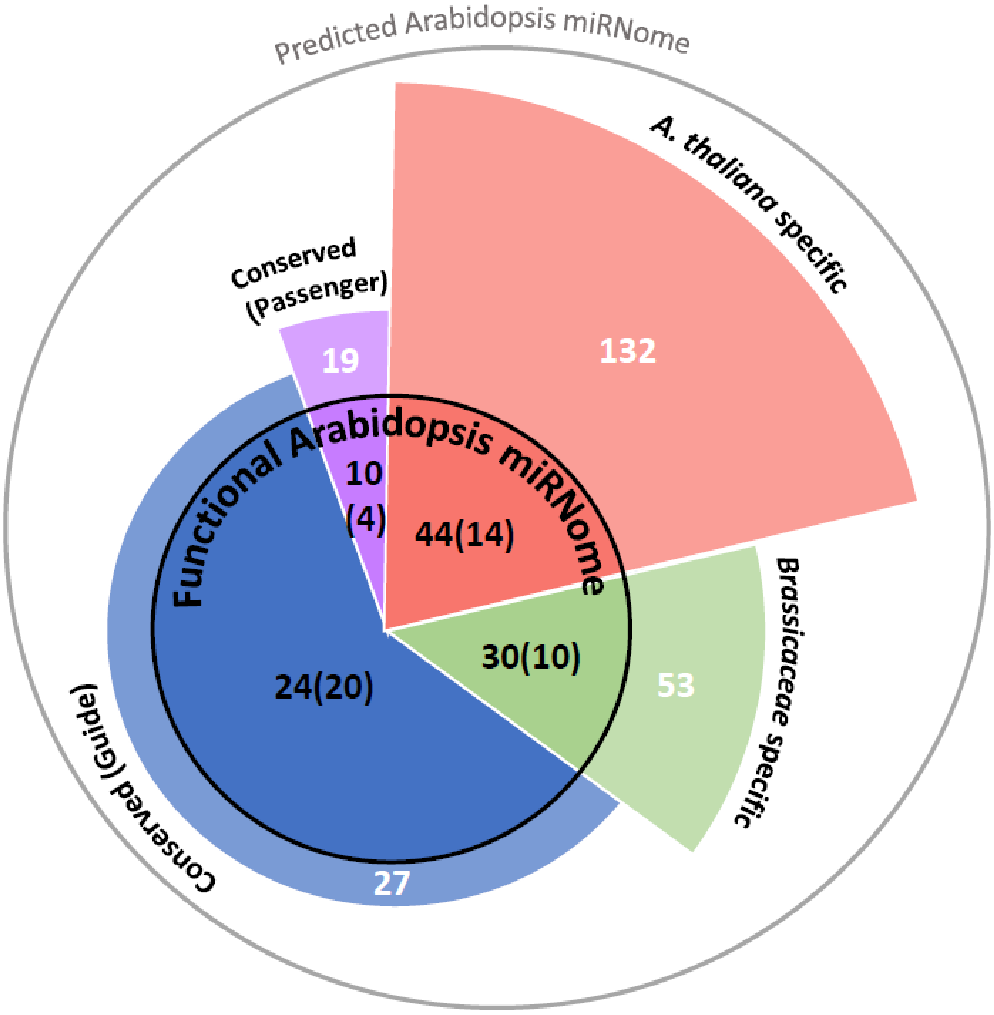


Figure 4.1. A proposed model of the functional Arabidopsis miRNome.
The outer grey circle represents the Arabidopsis miRNome as inferred by all entries from miRBase v22 (Kozomara et al., 2019). The white numbers indicate the number of miRNA families in each conservation group. The black circle represents the miRNome as inferred by TRUEE. The black numbers indicate the number of miRNA families identified with a HE target in each conservation group under low stringency and high stringency (bracketed) filters. The large discrepancy between the outer grey circle from the inner black circle suggests that the functional Arabidopsis miRNome to be much smaller than what is currently annotated. Although most miRNA families from miRBase v22 are predominantly from the *A. thaliana* specific and *Brassicaceae* specific categories, the Arabidopsis targetome is predominantly regulated by the conserved guide strands. This shows miRNA entries to be predominantly from the *A. thaliana*-specific then, *Brassicaceae*-specific categories. These were then followed by the conserved passenger and guide strand miRNAs.

A question arises as to why these conserved MTIs predominate the targetome. Their high conservation across distant plant clades may be attributed to the important biological processes the targets are involved in, such as, plant development, morphology, and stress response, resulting in these MTIs to undergo strong selection pressures. Many miRNA targets are transcription factors which regulate numerous downstream genes, and are core for plant biology (Reviewed in Jones-Rhoades et al., 2006). As retention of these conserved targets is important, this will restrict which additional targets can be acquired as they will also need to conform to the regulatory conditions defined by the former. This is because the expression pattern of the miRNA will be constrained by the required regulatory outcome of the predominating target family, and so the regulation of any additional targets must be achieved in the context of this miRNA expression pattern. Supporting this notion, is that many additional target families identified by TRUEE and in other studies are found to be involved in related pathways to the primary conserved miRNA target family (Chorostecki et al., 2012; Debernardi et al., 2012) (discussed in Chapter 3). The additional target families were also less conserved which may reflect that they have less functional importance than the predominant conserved target families and may be easily lost. Therefore, it could be generalized that there appears that there are three categories for the MTIs which make up the plant targetome. Firstly, conserved (fixed) MTIs which are involved in fundamental plant biology processes that are absolutely essential across land plants (e.g. leaf polarity, flowering, phase change) (eg. miR156:*SPL*, miR160:*ARF*; miR393:*AFB*). Secondly, less conserved (more fluid) MTIs which provide a specialized adaptive trait within a narrow range of plant species, but that are non-critical in other plant species (eg. miR396:*bHLH74*; miR827:*NLA*). Thirdly, transient young “miRNAs” which have yet to acquire a target interaction of functional importance (no example by defintion) (Figure 4.2).

Only the conserved miRNAs across species and the Arabidopsis targetome was investigated in this thesis. As such, it would be of interest to investigate the miRNome and targetome of other plant species using TRUEE. Although our results strongly support a narrow functional scope of miRNA-mediated regulation in plants they do not rule out a different targetome landscape in other species. Furthermore, analysing the targetome of these species may identify multiple previously undocumented MTIs with high Cat Scores. Although few undocumented MTIs with strong Cat Scores were identified in Arabidopsis, many other species are less well studied and therefore may have greater potential for target discovery.

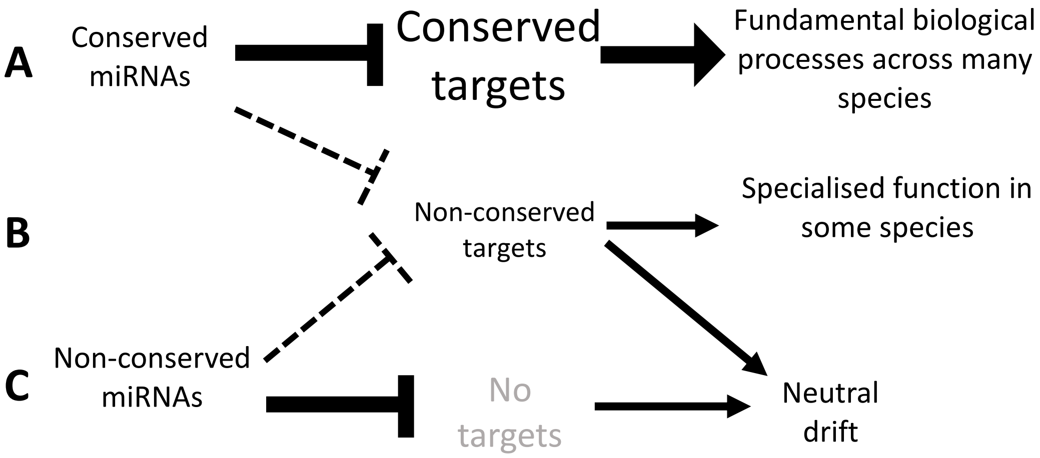


Figure 4.2. A proposed model of miRNA-mediated regulation in a plant targetome. A) Conserved miRNAs target many conserved homologous genes. These MTIs are under strong selective pressure and are core for fundamental plant biological processes. (eg. miR156:*SPL*, miR160:*ARF*; miR393:*AFB*) B) The acquisition of non-conserved targets by both conserved and non-conserved miRNAs is less common because it is constrained by the desired regulatory outcome of the conserved MTIs. These MTIs are more fluid as they depend on the specialized adaptive needs of the plant and may be lost over evolutionary time. (eg. miR396:*bHLH74*; miR827:*NLA*) C) Most non-conserved miRNAs do not acquire targets of functional significance and therefore undergo neutral drift. Only few acquire targets and most of these MTIs are also weaker and of less functional importance.

4.3 miRNA regulatory constraints and their implications to miRNA-based biotechnology and target prediction

Our results in Chapter 3 support a previous notion proposing there to be regulatory constraints which limit the scope of miRNA-mediated regulation on the plant transcriptome (Li et al., 2014). Firstly, we find greater evidence for factors beyond binding site complementarity which influence miRNA-mediated regulation. Secondly, as mentioned before, we find evidence supporting that the acquisition of targets is limited by its biological function. That is where its desired regulatory outcome must be compatible with that of the predominant conserved target family. Identifying these factors which are involved in MTIs is of great interests to improve miRNA target prediction programs and miRNA-based biotechnology, such as artificial miRNAs (amiRNAs), artificial small interfering RNAs (siRNAs) and miRNA decoys. Factors beyond miRNA-target binding site complementarity have been investigated in both plants and animal (Kertesz et al., 2007; Kedde et al., 2007; Kedde et al., 2010; Tan et al., 2011; Gu et al., 2012; Li et al., 2012; Zheng et al., 2017; Yang et al., 2020; Kakumani et al., 2021). However, to date, the understanding of these mechanisms is limited and individual examples of these remain relatively few and therefore, currently cannot be generalized as features of a MTI. For example, attempts made to apply features derived from animal studies in plant target prediction found that results were not reflective of the strength of the MTI (Li et al., 2013; Deveson et al., 2013; Zheng et al., 2017). Attempting to rank likely targets by binding site complementarity also did not correlate with a strong MTI (Liu et al., 2014; Dai et al., 2018; Millar et al., 2019).

An error in this approach may be that it assumes that, for these features, there is one-hard-fast set of rules that can be generally applied across all miRNAs and their targets. Rather, given the evolutionary history of these ancient MTIs, it is likely that plant miRNA-target family modules have arisen independently and evolved a unique combination of features to achieve a functionally relevant regulatory outcome. An example is the conserved sequences and secondary structures which are only found in the strongly regulated *MYB* targets of miR159, *MYB33/65* (Zheng et al., 2017). An explanation may be the different biological function *MYB33/65* has from all the other predicted *MYB* targets. Although, all the predicted *MYB* targets are involved in male organ development, only *MYB33/65* is transcribed widely across tissues (Millar & Gubler, 2005; Liang et al, 2013). *MYB33/65* expression is then restricted again to anthers by miR159-mediated regulation which is constitutively and strongly expressed. A potential function of this seemingly redundant miR159:*MYB33/65* expression pattern was recently proposed in a study by Zheng et al. (2020) in *N. tabacum*. It may be that *MYB33/65* is only released from miR159-mediated regulation across tissues for expression in response to pathogens. Therefore, it may be that these secondary structures have likely arisen independently to achieve this specific biological function unique to *MYB33/65* and cannot be widely applied across genes. Illustrating this, engineering an analogous secondary structure next to the miR159 binding site of the poorly regulated *MYB81* did not improve the regulation of this gene (Zheng, 2018). That features have arisen independently may also explain why in Chapter 3, although there appeared to be commonalities, features varied between different miRNA-target pairs. Conserved sequences were only found for some MTIs and not all, and the predicted RNA secondary structures of each MTI with conserved sequences were unique in structure.

That conserved MTIs have independently evolved unique combinations of features is also supported by our results finding that each miRNA-target family pair differed in the degree of miRNA-target binding site complementarity (Figure 3.5). We found that for several miRNA-target family modules, a high degree of complementarity correlated with HE targets [miR160*:ARF*; miR166:*HD-ZIPIII*; miR171:*HAM*; miR172:*AP2*; tasiARF:*ARF*]. Alternatively, other modules had lower complementarity requirements [miR159:*MYB*; miR169:*NF-YA*; miR319:*TCP*; miR395:*APS*]. For most miRNA-target family pairs, why the complementarity requirements vary is unknown to date, however, some mismatches appear conserved. For miR159:*MYB*, mismatches to the binding site at nt positions 1, 15-16 and 21 (corresponding to the miR159) were consistently found across diverse species (Millar et al., 2019). For miR396:*GRF*, the increased mismatch score is due to a 1 nt bulge created in the GRF binding site by miR396 which is conserved across GRF homologues and across diverse species (Debernardi et al., 2012; Bazin et al., 2013; Zhang et al., 2015). Although unknown in other species, the functional significance of this mismatch has been investigated for one Arabidopsis homologue. A study by Debernardi et al. (2012) reported that these mismatches were required for the desired regulatory outcome of both the *GRF* family member, *GRF2*, and a secondary target, *bHLH72*, for proper leaf development. This mismatch reduces the strength of silencing on *GRF2* and enables the regulation of *bHLH72*. In this case, it appears that not only are mismatches tolerated, but are selected for. Similarly, for miR159:*MYB*, mismatches to the binding site at nt positions 1, 15-16 and 21 (corresponding to the miR159) were consistently found across diverse species (Millar et al., 2019) but the reason for this conservation is unknown. As such, although miRNA-target binding site complementarity is necessary, it is not an absolute indicator of an MTI in plants.

That miRNA-target modules have co-evolved unique features to satisfy a desired regulatory outcome has implications for miRNA-based biotechnology. This can be seen in that despite being designed with high complementarity to their intended targets, off-target effects of siRNAs and amiRNAs remains a commonly reported problem (Xu et al., 2006; Li et al., 2013; Deveson et al., 2013). Rather, it has been previously proposed that natural miRNAs have co-evolved with the rest of the transcriptome to avoid promiscuous targeting (Bartel & Chen, 2004; Schwab et al., 2005). As such, there may be no one hard-and-fast rule in designing an amiRNA or siRNA. This may also explain why different miRNA decoys designed to target the same miRNA, and conversely, the same miRNA decoy designed to target different miRNAs perform with varying efficacies (Todesco et al., 2010; Ivashuta et al., 2011; Reichel et al., 2015). The implication of these results suggests that we may be approaching a limit in identifying general rules which can be applied to improve target prediction programs and programs that guide the design of amiRNAs, siRNAs and miRNA decoys. Rather, designing and experimentally testing multiple amiRNAs, siRNA or miRNA decoys to find one with the desired regulatory outcome may be necessary. An improvement resulting from this thesis may be the addition of a filter to miRNA-target prediction programs to prioritise genes that are functionally related to the predominant miRNA target family for further investigation for miRNA-mediated regulation. This may be achieved using TRUEE followed by analysis to identify which targets are related in molecular function and biochemical pathways, and/or share an overlapping expression pattern.

4.4 Investigation of the conserved sequences flanking the miRNA target binding sites

Only a few features beyond binding site complementarity involved in miRNA-mediated regulation have been investigated in plants (Zheng et al., 2017; Yang et al., 2020). This thesis found further examples of these features in plants. Multiple miRNA-target family pairs possessed conserved sequences which appeared to be correlated with miRNA-mediated regulation.

Furthermore, many of the conserved sequence miRNA-target pairs were predicted to form secondary structure. Whether these conserved sequences or RNA secondary structures impact miRNA-mediated regulation requires further experimentation to determine. The study by Zheng et al. (2017) on miR159:*MYB33* implicated two stem-loops flanking the miR159 binding site in miRNA-mediated regulation in a functional genetic study *in vivo*. Introducing 6 or 7 nt synonymous mutations that were predicted to destroy either of the stem-loops disrupted miRNA-mediated regulation, but restoring one of the stem-loops with a further 6 nt synonymous mutations restored regulation. However, it may not be possible to use this approach for *ARF10* and some of the other targets due to a much smaller number of possible sites for mutation without changing the amino acid sequence. Furthermore, even using this approach, the secondary structures still remain a prediction. These are determined by *in silico* algorithms calculating the minimum free energy for the most thermodynamically stable structure (Bernhart et al., 2014). Furthermore, RNA secondary structure has been reported to be dynamic and so using the most thermodynamically stable structure may not be reflective of its conformation in an *in vivo* environment (Rouskin et al., 2014). These conformations may change under different cellular conditions to achieve a desired regulatory outcome of the target gene. One such example in animals is the conformational change in the miR-221 target, *p27*, caused by the RNA binding protein, Pumilio1, binding proximally to the binding site (Kedde et al., 2010). Furthermore, the secondary structure may differ with subtle changes to the RNA sequence and with using different algorithmic models used which are an available option in bioinformatic prediction programs (Lorenz et al., 2011; Bernhart et al., 2014).

Therefore, investigating if these RNA secondary structures form using *in vivo* experimental evidence is required. One such method is dimethyl sulphate sequencing (DMS) which enables a transcriptome-wide analysis of RNA secondary structures *in vivo* in plants (Ding et al., 2014). DMS methylates the N1 of adenosine and the N3 of cytosine in unstructured RNA regions such as the loops regions of stem-loops, mismatches and bulges and therefore allows the RNA secondary structure to be inferred. However, a potential caveat to this approach may be the difficulty in capturing target gene transcripts as their strong silencing would presumably lead to low transcript levels. Overall, implicating whether a secondary structure proximal to a miRNA-binding site is involved in miRNA-mediated regulation still remains a challenge.

References

Addo-Quaye, C., Miller, W., & Axtell, M. J. (2009). CleaveLand: A pipeline for using degradome data to find cleaved small RNA targets. *Bioinformatics*. https://doi.org/10.1093/bioinformatics/btn604

Axtell, M. J. (2008). Evolution of microRNAs and their targets: Are all microRNAs biologically relevant? *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, *1779*(11), 725–734. https://doi.org/10.1016/j.bbagrm.2008.02.007

Axtell, M. J., & Meyers, B. C. (2018). Revisiting Criteria for Plant MicroRNA Annotation in the Era of Big Data. *The Plant Cell*, *30*(2), 272–284. https://doi.org/10.1105/tpc.17.00851

Axtell, M. J., Snyder, J. A., & Bartel, D. P. (2007). Common Functions for Diverse Small RNAs of Land Plants. *The Plant Cell*, *19*(6), 1750–1769. https://doi.org/10.1105/tpc.107.051706

Bartel, D. P., & Chen, C.-Z. (2004). Micromanagers of gene expression: the potentially widespread influence of metazoan microRNAs. *Nature Reviews Genetics*, *5*(5), 396–400. https://doi.org/10.1038/nrg1328

Bazin, J., Khan, G. A., Combier, J.-P., Bustos-Sanmamed, P., Debernardi, J. M., Rodriguez, R., Sorin, C., Palatnik, J., Hartmann, C., Crespi, M., & Lelandais-Brière, C. (2013). miR396 affects mycorrhization and root meristem activity in the legume Medicago truncatula. *The Plant Journal*, *74*(6), 920–934. https://doi.org/https://doi.org/10.1111/tpj.12178

Bernhart, S. H., Hofacker, I. L., Will, S., Gruber, A. R., & Stadler, P. F. (2008). RNAalifold: improved consensus structure prediction for RNA alignments. *BMC Bioinformatics*, *9*(1), 474. https://doi.org/10.1186/1471-2105-9-474

Brousse, C., Liu, Q., Beauclair, L., Deremetz, A., Axtell, M. J., & Bouché, N. (2014). A non-canonical plant microRNA target site. *Nucleic Acids Research*, *42*(8), 5270–5279. https://doi.org/10.1093/nar/gku157

Bülow, L., Bolívar, J. C., Ruhe, J., Brill, Y., & Hehl, R. (2012). “MicroRNA Targets”, a new AthaMap web-tool for genome-wide identification of miRNA targets in Arabidopsis thaliana. *BioData Mining*, *5*(1), 7. https://doi.org/10.1186/1756-0381-5-7

Chávez Montes, R. A., De Fátima Rosas-Cárdenas, F., De Paoli, E., Accerbi, M., Rymarquis, L. A., Mahalingam, G., Marsch-Martínez, N., Meyers, B. C., Green, P. J., & De Folter, S. (2014). Sample sequencing of vascular plants demonstrates widespread conservation and divergence of microRNAs. *Nature Communications*. https://doi.org/10.1038/ncomms4722

Chorostecki, U., Crosa, V. A., Lodeyro, A. F., Bologna, N. G., Martin, A. P., Carrillo, N., Schommer, C., & Palatnik, J. F. (2012). Identification of new microRNA-regulated genes by conserved targeting in plant species. *Nucleic Acids Research*, *40*(18), 8893–8904. https://doi.org/10.1093/nar/gks625

Cui, J., You, C., & Chen, X. (2017). The evolution of microRNAs in plants. In *Current Opinion in Plant Biology*. https://doi.org/10.1016/j.pbi.2016.11.006

Cuperus, J. T., Fahlgren, N., & Carrington, J. C. (2011). Evolution and functional diversification of MIRNA genes. *The Plant Cell*, *23*(2), 431–442. https://doi.org/10.1105/tpc.110.082784

Dai, X., Zhuang, Z., & Zhao, P. X. (2018). PsRNATarget: A plant small RNA target analysis server (2017 release). *Nucleic Acids Research*, *46*(W1), W49–W54. https://doi.org/10.1093/nar/gky316

Debernardi, J. M., Rodriguez, R. E., Mecchia, M. A., & Palatnik, J. F. (2012). Functional specialization of the plant miR396 regulatory network through distinct microRNA-target interactions. *PLoS Genetics*. https://doi.org/10.1371/journal.pgen.1002419

Ding, Y., Tang, Y., Kwok, C. K., Zhang, Y., Bevilacqua, P. C., & Assmann, S. M. (2014). In vivo genome-wide profiling of RNA secondary structure reveals novel regulatory features. *Nature*, *505*(7485), 696–700. https://doi.org/10.1038/nature12756

Fahlgren, N., Howell, M. D., Kasschau, K. D., Chapman, E. J., Sullivan, C. M., Cumbie, J. S., Givan, S. A., Law, T. F., Grant, S. R., Dangl, J. L., & Carrington, J. C. (2007). High-throughput sequencing of Arabidopsis microRNAs: Evidence for frequent birth and death of MIRNA genes. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0000219

Fei, Y., Mao, Y., Shen, C., Wang, R., Zhang, H., & Huang, J. (2020). WPMIAS: Whole-degradome-based Plant MicroRNA–target Interaction Analysis Server. *Bioinformatics*, *36*(6), 1937–1939. https://doi.org/10.1093/bioinformatics/btz820

Folkes, L., Moxon, S., Woolfenden, H. C., Stocks, M. B., Szittya, G., Dalmay, T., & Moulton, V. (2012). PAREsnip: A tool for rapid genome-wide discovery of small RNA/target interactions evidenced through degradome sequencing. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gks277

Gu, W., Wang, X., Zhai, C., Xie, X., & Zhou, T. (2012). Selection on synonymous sites for increased accessibility around mirna binding sites in plants. *Molecular Biology and Evolution*, *29*(10), 3037–3044. https://doi.org/10.1093/molbev/mss109

Ivashuta, S., Banks, I. R., Wiggins, B. E., Zhang, Y., Ziegler, T. E., Roberts, J. K., & Heck, G. R. (2011). Regulation of gene expression in plants through miRNA inactivation. *PLoS ONE*, *6*(6). https://doi.org/10.1371/journal.pone.0021330

Jones-Rhoades, M. W., Bartel, D. P., & Bartel, B. (2006). MicroRNAs and Their Regulatory Roles in Plants. *Plant Biology*, 1–16. https://doi.org/10.1146/annurev.arplant.57.032905.105218

Kakumani, P. K., Guitart, T., Houle, F., Harvey, L.-M., Goyer, B., Germain, L., Gebauer, F., & Simard, M. J. (2021). CSDE1 attenuates microRNA-mediated silencing of PMEPA1 in melanoma. *Oncogene*, *40*(18), 3231–3244. https://doi.org/10.1038/s41388-021-01767-9

Kedde, M., Strasser, M. J., Boldajipour, B., Vrielink, J. A. F. O., Slanchev, K., le Sage, C., Nagel, R., Voorhoeve, P. M., van Duijse, J., Ørom, U. A., Lund, A. H., Perrakis, A., Raz, E., & Agami, R. (2007). RNA-Binding Protein Dnd1 Inhibits MicroRNA Access to Target mRNA. *Cell*, *131*(7), 1273–1286. https://doi.org/https://doi.org/10.1016/j.cell.2007.11.034

Kedde, M., van Kouwenhove, M., Zwart, W., Oude Vrielink, J. A. F., Elkon, R., & Agami, R. (2010). A Pumilio-induced RNA structure switch in p27-3’ UTR controls miR-221 and miR-

222 accessibility. *Nature Cell Biology*, *12*(10), 1014–1020. https://doi.org/10.1038/ncb2105

Kertesz, M., Iovino, N., Unnerstall, U., Gaul, U., & Segal, E. (2007). The role of site accessibility in microRNA target recognition. *Nature Genetics*, *39*(10), 1278–1284. https://doi.org/10.1038/ng2135

Kozomara, A., Birgaoanu, M., & Griffiths-jones, S. (2019). *miRBase : from microRNA sequences to function*. *47*(November 2018), 155–162. https://doi.org/10.1093/nar/gky1141

Li, F., Zheng, Q., Vandivier, L. E., Willmann, M. R., Chen, Y., & Gregory, B. D. (2012). Regulatory impact of RNA secondary structure across the Arabidopsis transcriptome. *The Plant Cell*, *24*(11), 4346–4359. https://doi.org/10.1105/tpc.112.104232

Li, J.-F., Chung, H. S., Niu, Y., Bush, J., McCormack, M., & Sheen, J. (2013). Comprehensive protein-based artificial microRNA screens for effective gene silencing in plants. *The Plant Cell*, *25*(5), 1507–1522. https://doi.org/10.1105/tpc.113.112235

Li, J., Reichel, M., Li, Y., & Millar, A. A. (2014). The functional scope of plant microRNA-mediated silencing. *Trends in Plant Science*, *19*(12), 750–756. https://doi.org/10.1016/j.tplants.2014.08.006

Liang, Y., Tan, Z.-M., Zhu, L., Niu, Q.-K., Zhou, J.-J., Li, M., Chen, L.-Q., Zhang, X.-Q., & Ye, D. (2013). MYB97, MYB101 and MYB120 Function as Male Factors That Control Pollen Tube-Synergid Interaction in Arabidopsis thaliana Fertilization. *PLOS Genetics*, *9*(11), e1003933. https://doi.org/10.1371/journal.pgen.1003933

Lindow, M., Jacobsen, A., Nygaard, S., Mang, Y., & Krogh, A. (2007). Intragenomic Matching Reveals a Huge Potential for miRNA-Mediated Regulation in Plants. *PLOS Computational Biology*, *3*(11), e238. https://doi.org/10.1371/journal.pcbi.0030238

Lindow, M., & Krogh, A. (2005). Computational evidence for hundreds of non-conserved plant microRNAs. *BMC Genomics*, *6*(1), 119. https://doi.org/10.1186/1471-2164-6-119

Liu, Q., Wang, F., & Axtell, M. J. (2014). Analysis of Complementarity Requirements for Plant MicroRNA Targeting Using a Nicotiana benthamiana Quantitative Transient Assay . *The Plant Cell*, *26*(2), 741–753. https://doi.org/10.1105/tpc.113.120972

Liu, W. W., Meng, J., Cui, J., & Luan, Y. S. (2017). Characterization and function of MicroRNA∗s in plants. *Frontiers in Plant Science*, *8*(December), 1–7. https://doi.org/10.3389/fpls.2017.02200

Lorenz, R., Bernhart, S. H., Höner zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P. F., & Hofacker, I. L. (2011). ViennaRNA Package 2.0. *Algorithms for Molecular Biology*, *6*(1), 26. https://doi.org/10.1186/1748-7188-6-26

Meng, Y., Shao, C., Gou, L., Jin, Y., & Chen, M. (2011). Construction of microRNA- and microRNA*- mediated regulatory networks in plants. *RNA Biology*, *8*(6). https://doi.org/10.4161/rna.8.6.17743

Meng, Y., Shao, C., Wang, H., & Chen, M. (2012). Are all the miRBase-registered microRNAs true? *RNA Biology*. https://doi.org/10.4161/rna.19230

Millar, A. A., & Gubler, F. (2005). The Arabidopsis GAMYB-Like Genes, MYB33 and MYB65, Are MicroRNA-Regulated Genes That Redundantly Facilitate Anther Development. *The Plant Cell*, *17*(March), 705–721. https://doi.org/10.1105/tpc.104.027920.)

Millar, A. A., Lohe, A., & Wong, G. (2019). Biology and Function of miR159 in Plants. *Plants (Basel, Switzerland)*, *8*(8), 255. https://doi.org/10.3390/plants8080255

Rajagopalan, R., Vaucheret, H., Trejo, J., & Bartel, D. P. (2006). A diverse and evolutionarily fluid set of microRNAs in Arabidopsis thaliana. *Genes and Development*, *20*(24), 3407–3425. https://doi.org/10.1101/gad.1476406

Reichel, M., Li, Y., Li, J., & Millar, A. A. (2015). Inhibiting plant microRNA activity: Molecular SPONGEs, target MIMICs and STTMs all display variable efficacies against target microRNAs. *Plant Biotechnology Journal*, 1–12. https://doi.org/10.1111/pbi.12327

Rouskin, S., Zubradt, M., Washietl, S., Kellis, M., & Weissman, J. S. (2014). Genome-wide probing of RNA structure reveals active unfolding of mRNA structures in vivo. *Nature*, *505*(7485), 701–705. https://doi.org/10.1038/nature12894

Schwab, R., Palatnik, J. F., Riester, M., Schommer, C., Schmid, M., & Weigel, D. (2005). Specific effects of microRNAs on the plant transcriptome. *Developmental Cell*, *8*(4), 517–527. https://doi.org/10.1016/j.devcel.2005.01.018

Smith, L. M., Burbano, H. A., Wang, X., Fitz, J., Wang, G., Ural-Blimke, Y., & Weigel, D. (2015). Rapid divergence and high diversity of miRNAs and miRNA targets in the Camelineae. *Plant Journal*. https://doi.org/10.1111/tpj.12754

Tay, Y., Kats, L., Salmena, L., Weiss, D., Tan, S. M., Ala, U., Karreth, F., Poliseno, L., Provero, P., Di Cunto, F., Lieberman, J., Rigoutsos, I., & Pandolfi, P. P. (2011). Coding-independent regulation of the tumor suppressor PTEN by competing endogenous mRNAs. *Cell*, *147*(2), 344–357. https://doi.org/10.1016/j.cell.2011.09.029

Taylor, R. S., Tarver, J. E., Foroozani, A., & Donoghue, P. C. J. (2017). *Insights & Perspectives MicroRNA annotation of plant genomes À Do it right or not at all*. *1600113*, 1–6. https://doi.org/10.1002/bies.201600113

Taylor, R. S., Tarver, J. E., Hiscock, S. J., & Donoghue, P. C. J. (2014). Evolutionary history of plant microRNAs. *Trends in Plant Science*, *19*(3), 175–182. https://doi.org/10.1016/j.tplants.2013.11.008

Todesco, M., Rubio-Somoza, I., Paz-Ares, J., & Weigel, D. (2010). A collection of target mimics for comprehensive analysis of MicroRNA function in Arabidopsis thaliana. *PLoS Genetics*, *6*(7), 1–10. https://doi.org/10.1371/journal.pgen.1001031

Xu, P., Zhang, Y., Kang, L., Roossinck, M. J., & Mysore, K. S. (2006). Computational Estimation and Experimental Verification of Off-Target Silencing during Posttranscriptional Gene Silencing in Plants. *Plant Physiology*, *142*(2), 429–440. https://doi.org/10.1104/pp.106.083295

Yang, M., Woolfenden, H. C., Zhang, Y., Fang, X., Liu, Q., Vigh, M. L., Cheema, J., Yang, X., Norris, M., Yu, S., Carbonell, A., Brodersen, P., Wang, J., & Ding, Y. (2020). Intact RNA structurome reveals mRNA structure-mediated regulation of miRNA cleavage in vivo. *Nucleic Acids Research*, *48*(15), 8767–8781. https://doi.org/10.1093/nar/gkaa577

Zhang, K., Shi, X., Zhao, X., Ding, D., Tang, J., & Niu, J. (2015). Investigation of miR396 and growth-regulating factor regulatory network in maize grain filling. *Acta Physiologiae Plantarum*, *37*(2), 28. https://doi.org/10.1007/s11738-014-1767-6

Zheng, Z., Reichel, M., Deveson, I., Wong, G., Li, J., & Millar, A. A. (2017). Target RNA Secondary Structure Is a Major Determinant of miR159 Efficacy. *Plant Physiology*, *174*(3), 1764–1778. https://doi.org/10.1104/pp.16.01898

Zheng, Z. (2018). The miR159-GAMYB pathway: silencing and function of GAMYB homologues amongst diverse plant species [Doctoral thesis, Australian National University]. ANU Open Research Repository. https://openresearch-repository.anu.edu.au/bitstream/1885/148731/1/Zheng%20Thesis%202018.pdf

Zheng, Z., Wang, N., Jalajakumari, M., Blackman, L., Shen, E., Verma, S., Wang, M. B., & Millar, A. A. (2020). MiR159 represses a constitutive pathogen defense response in tobacco. *Plant Physiology*, *182*(4), 2182–2198. https://doi.org/10.1104/PP.19.00786

Appendix

The Supplementary Tables and Figures for Chapter 2 and Chapter 3 are accessible on the Open Science Framework page **https://osf.io/s83t6/**.

Figure S1. T-plots of HE targets not from the VAT set found at a Library % Cut-off of 40%

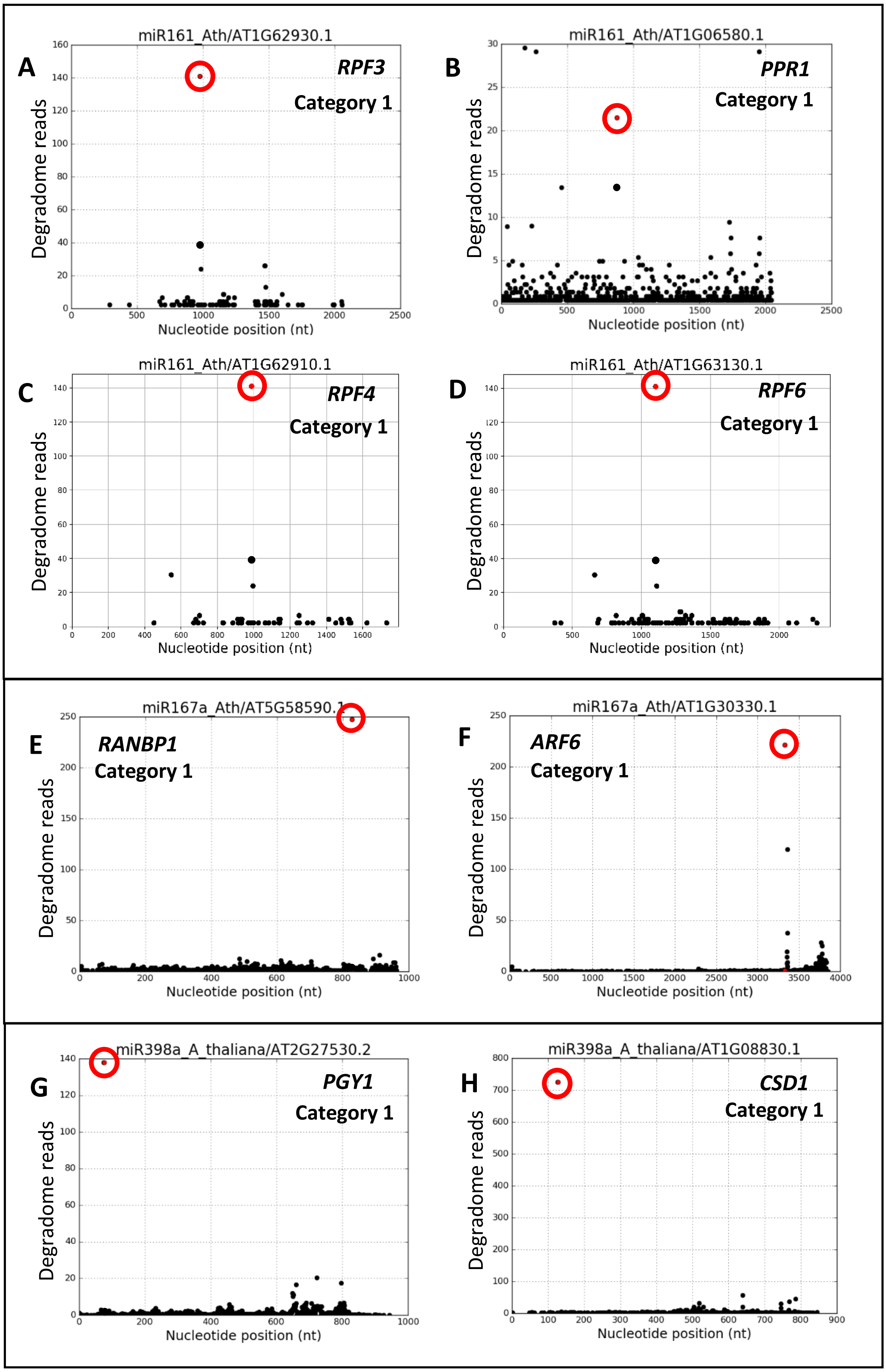
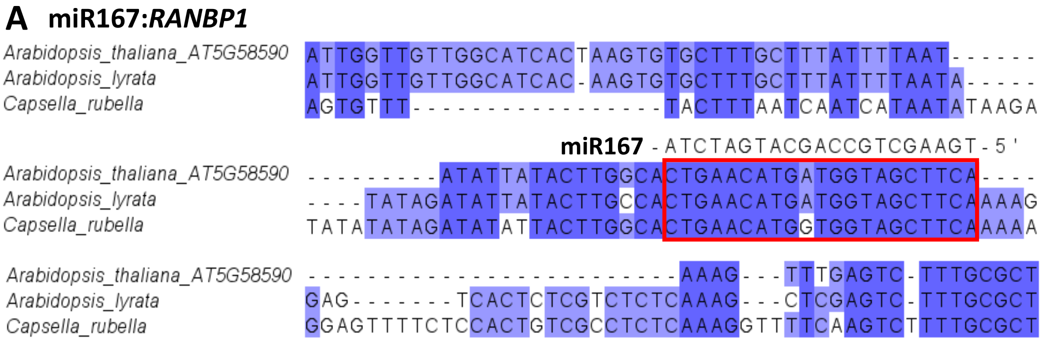


Figure S1. T-plots of HE targets not from the VAT set found at a Library % Cut-off of 40%. T-Plots of (A) *RNA PROCESSING FACTOR 3* (*RPF3*); (B) *PENTATRICOPEPTIDE REPEAT 1* (*PPR1*); (C) *RPF4* and (D) *RPF6*; (E) *RNA BINDING PROTEIN 1* (*RANBP1*); (F) *AUXIN RESPONSE FACTOR 6* (*ARF6*);

(G) *PIGGYBACK1* (*PGY1*) that encodes a ribosomal protein L10aP; (H) *COPPER/ZINC SUPEROXIDE DISMUTASE 1* (*CSD1*). The T-plot from the degradome library with the highest *Maximum Category* and highest *Cleavage Tag Abundance* was used for each miRNA target. The cleavage tag is circled in red. T-plot figures were adapted from WPMIAS (Fei et al., 2020).

Figure S2. Binding site conservation of HE targets is limited to the *Brassicaceae* family

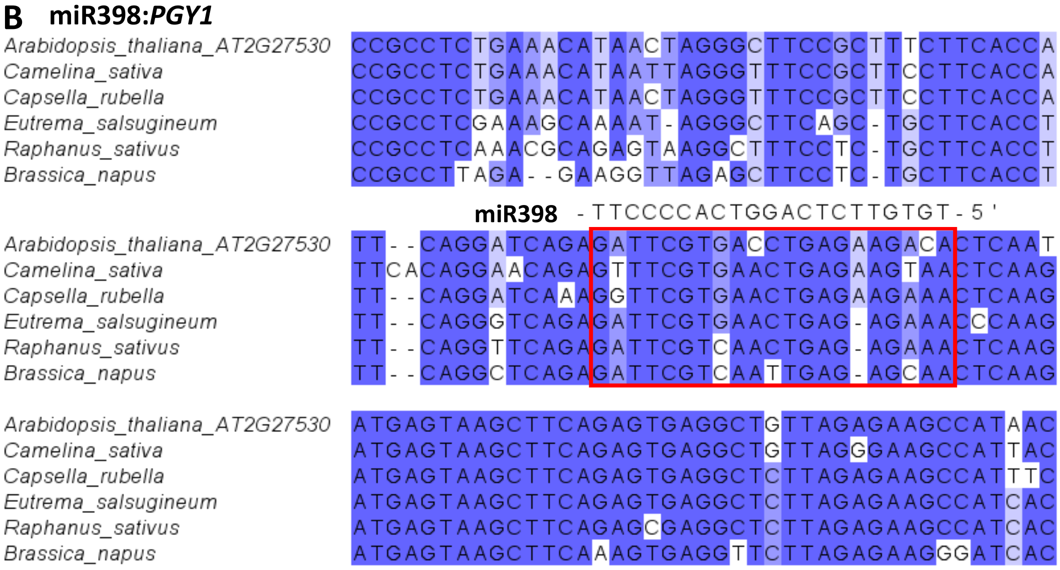


A **miR167:*RANBP1***

```
Arabidopsis_thaliana_AT5G58590  ATTGGTTGTTGGCATCACTAAGTGTGCTTTGCTTTATTTTAAT------
Arabidopsis_lyrata              ATTGGTTGTTGGCATCAC-AAGTGTGCTTTGCTTTATTTTAATA-----
Capsella_rubella                AGTGTTT-----------------TACTTTAATCAATCATAATATAAGA

                                         miR167 -ATCTAGTACGACCGTCGAAGT-5'
Arabidopsis_thaliana_AT5G58590  ---------ATATTATACTTGGCACTGAACATGATGGTAGCTTCA----
Arabidopsis_lyrata              ----TATAGATATTATACTTGCCACTGAACATGATGGTAGCTTCAAAAG
Capsella_rubella                TATATATAGATATATTACTTGGCACTGAACATGGTGGTAGCTTCAAAAA

Arabidopsis_thaliana_AT5G58590  -------------------------AAAG---TTTGAGTC-TTTGCGCT
Arabidopsis_lyrata              GAG-------TCACTCTCGTCTCTCAAAG---CTCGAGTC-TTTGCGCT
Capsella_rubella                GGAGTTTTCTCCACTGTCGCCTCTCAAAGGTTTTCAAGTCTTTTGCGCT
```



B **miR398:*PGY1***

```
Arabidopsis_thaliana_AT2G27530  CCGCCTCTGAAACATAACTAGGGCTTCCGCTTTCTTCACCA
Camelina_sativa                 CCGCCTCTGAAACATAATTAGGGTTTCCGCTTCCTTCACCA
Capsella_rubella                CCGCCTCTGAAACATAACTAGGGTTTCCGCTTCCTTCACCA
Eutrema_salsugineum             CCGCCTCGAAAGCAAAAT-AGGGCTTCAGC-TGCTTCACCT
Raphanus_sativus                CCGCCTCAAACGCAGAGTAAGGCTTTCCTC-TGCTTCACCT
Brassica_napus                  CCGCCTTAGA--GAAGGTTAGAGCTTCCTC-TGCTTCACCT

                                  miR398 -TTCCCCACTGGACTCTTGTGT-5'
Arabidopsis_thaliana_AT2G27530  TT--CAGGATCAGAGATTCGTGACCTGAGAAGACACTCAAT
Camelina_sativa                 TTCACAGGAACAGAGTTTCGTGAACTGAGAAGTAACTCAAG
Capsella_rubella                TT--CAGGATCAAAGGTTCGTGAACTGAGAAGAAACTCAAG
Eutrema_salsugineum             TT--CAGGGTCAGAGATTCGTGAACTGAG-AGAAACCCAAG
Raphanus_sativus                TT--CAGGTTCAGAGATTCGTCAACTGAG-AGAAACTCAAG
Brassica_napus                  TT--CAGGCTCAGAGATTCGTCAATTGAG-AGCAACTCAAG

Arabidopsis_thaliana_AT2G27530  ATGAGTAAGCTTCAGAGTGAGGCTGTTAGAGAAGCCATAAC
Camelina_sativa                 ATGAGTAAGCTTCAGAGTGAGGCTGTTAGGGAAGCCATTAC
Capsella_rubella                ATGAGTAAGCTTCAGAGTGAGGCTCTTAGAGAAGCCATTTC
Eutrema_salsugineum             ATGAGTAAGCTTCAGAGTGAGGCTCTTAGAGAAGCCATCAC
Raphanus_sativus                ATGAGTAAGCTTCAGAGCGAGGCTCTTAGAGAAGCCATCAC
Brassica_napus                  ATGAGTAAGCTTCAAAGTGAGGTTCTTAGAGAAGGGATCAC
```

Figure S2. Binding site conservation of HE targets is limited to the *Brassicaceae* family. A multiple sequence alignment of A) miR167 HE target, *RANBP1*, and B) miR398 HE target, *PGY1*. The binding site is indicated in red with the miRNA sequence provided above. All species listed are from the *Brassicaceae* family.

Figure S3. Criteria required to determine the presence of conserved sequences

miRNA:Target family module	Criteria determining the presence of conserved sequences
miR159:MYB	• Must contain SL1 and/or SL2 (refer to Figure 3.7 A) • SL1 must contain sequences 'AGCAAT' followed by 'ATGGCA' separated by 9-11 nts as determined by MSA • SL2 must contain sequences 'CTTC' followed by 'GAAG' separated by 18-21 nts
miR160:ARF	• Must contain CS1 and/or CS2 (refer to Figure 3.21 A) • CS1 and CS2 must occur 8 nt immediately upstream and downstream of the binding site, respectively • CS1 is derived from the ARF10 and ARF17 conserved sequence • CS2 is derived from the ARF10 downstream sequence
miR171:HAM	• Must contain any conserved sequence (CS1-CS7) (refer to Figure 3.21 B) • CS1 must occur on the 30-35th nt upstream of the binding site • CS2 must occur on the 1-3rd nt upstream of the binding site • CS3 must occur on the 1-3rd nt downstream of the binding site • CS4 must occur on the 20-26th nt downstream of the binding site • CS5 must occur on the 32-38th nt downstream of the binding site • CS6 must occur on the 38-44th nt downstream of the binding site • CS7 must occur on the 44-50th nt downstream of the binding site
miR319:TCP	• Must contain any conserved sequence (CS1-CS3) (refer to Figure 3.21 C) • CS1 must occur on the 5-6th nt upstream of the binding site • CS2 must occur on the 1st nt upstream of the binding site • CS3 must occur on the 14th nt upstream of the binding site • CS1 is derived from the TCP2 conserved sequence • CS2 and CS3 are derived from the TCP4 conserved sequences
miR396:GRF	• Must contain any conserved sequence (CS1-CS4) (refer to Figure 3.21 D) • CS1 must occur on the 30th nt upstream of the binding site • CS2 must occur on the 9th nt upstream of the binding site • CS3 must occur on the 1st nt upstream of the binding site
TasiARF:ARF	• Must contain any conserved sequence (CS1-CS4) (refer to Figure 3.21 E) • CS1 must occur on the 20-22nd nt upstream of the binding site • CS2 must occur on the 13-15th nt upstream of the binding site • CS3 must occur on the 2nd nt upstream of the binding site • CS4 must occur on the 6th nt downstream of the binding site • CS1, CS2 and CS3 are derived from the conserved sequences of ARF2/3/4 • CS4 is derived from the downstream conserved sequence of ARF2

***All parameters are derived from the MSAs of target gene families**

Figure S4. Schematic of conserved sequences for all miRNA-target families flanking the binding sites

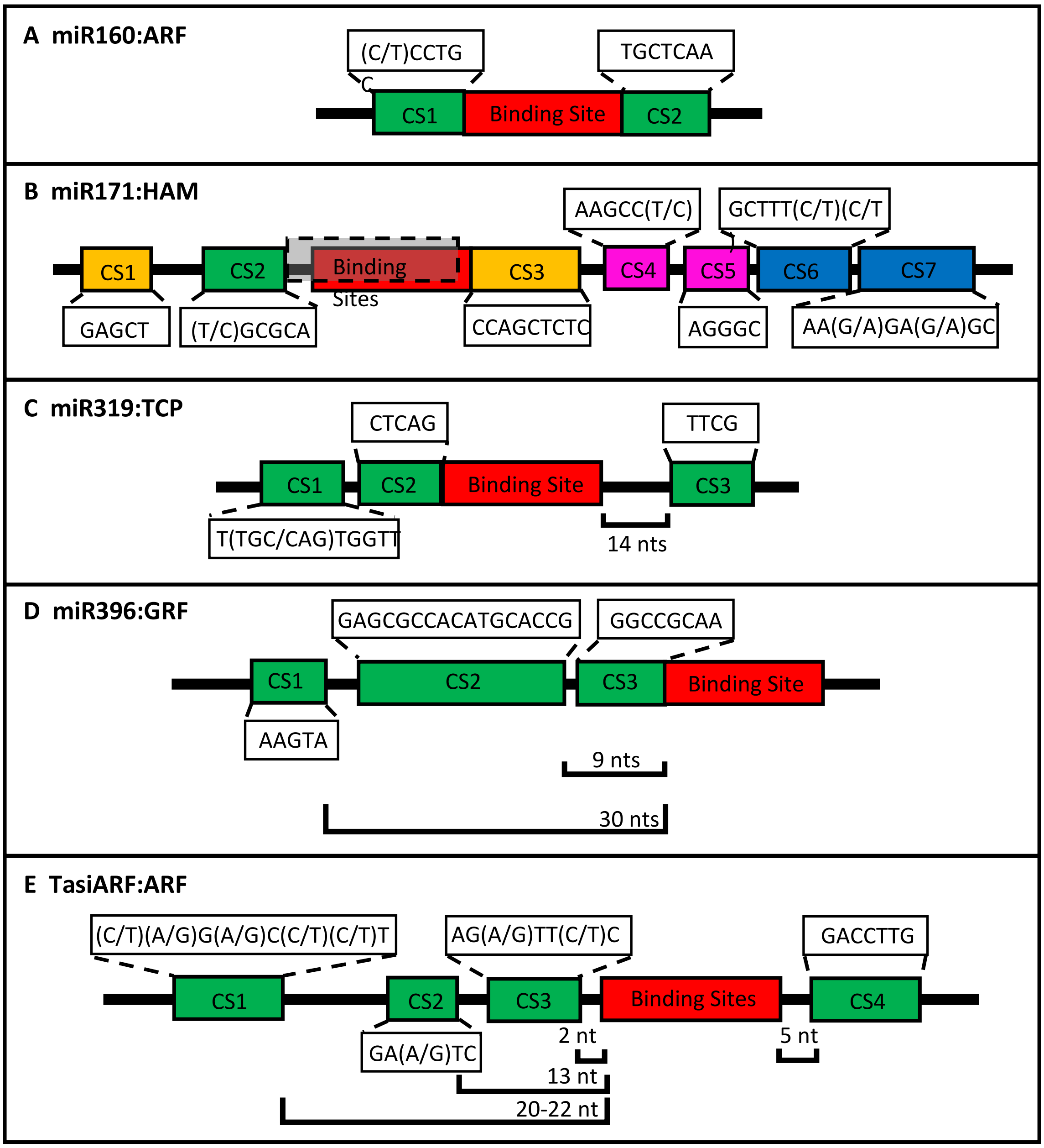


Figure S4. Schematic of conserved sequences for all miRNA-target families flanking the binding sites. Schematics of the conserved sequences flanking the binding site and their locations as determined by MSA for A) miR160:*ARF*, B) miR171:*HAM*, C) miR319:*TCP*, D) miR396:*GRF*, E) TasiARF:*ARF*. 'CS' stands for 'Conserved sequence' and the sequence is shown in white boxes. Nucleotides in parentheses indicate nucleotide variations due to compensatory nucleotide substitutions and alternate conserved sequences. The binding site is in red and additionally grey for miR171:*HAM1* as it has two overlapping binding sites.

References for Table S1

Abdel-Ghany, S. E., & Pilon, M. (2008). MicroRNA-mediated systemic down-regulation of copper protein expression in response to low copper availability in Arabidopsis. *Journal of Biological Chemistry*, *283*(23), 15932–15945. https://doi.org/10.1074/jbc.M801406200

Addo-Quaye, C., Eshoo, T. W., Bartel, D. P., & Axtell, M. J. (2008). Endogenous siRNA and miRNA Targets Identified by Sequencing of the Arabidopsis Degradome. *Current Biology*, *18*(10), 758–762. https://doi.org/10.1016/j.cub.2008.04.042

Allen, E., Xie, Z., Gustafson, A. M., & Carrington, J. C. (2005). microRNA-directed phasing during trans-acting siRNA biogenesis in plants. *Cell*, *121*(2), 207–221. https://doi.org/10.1016/j.cell.2005.04.004

Allen, E., Xie, Z., Gustafson, A. M., Sung, G. H., Spatafora, J. W., & Carrington, J. C. (2004). Evolution of microRNA genes by inverted duplication of target gene sequences in Arabidopsis thaliana. *Nature Genetics*, *36*(12), 1282–1290. https://doi.org/10.1038/ng1478

Allen, R. S., Li, J., Stahle, M. I., Dubroué, A., Gubler, F., & Millar, A. A. (2007). Genetic analysis reveals functional redundancy and the major target genes of the Arabidopsis miR159 family. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(41), 16371–16376. https://doi.org/10.1073/pnas.0707653104

Alonso-Peral, M. M., Li, J., Li, Y., Allen, R. S., Schnippenkoetter, W., Ohms, S., White, R. G., & Millar, A. A. (2010). The MicroRNA159-Regulated GAMYB-like Genes Inhibit Growth and Promote Programmed Cell Death in Arabidopsis. *Plant Physiol.*, *154*(2), 757–771. https://doi.org/10.1104/pp.110.160630

Alves-Junior, L., Niemeier, S., Hauenschild, A., Rehmsmeier, M., & Merkle, T. (2009). Comprehensive prediction of novel microRNA targets in Arabidopsis thaliana. *Nucleic Acids Research*, *37*(12), 4010–4021. https://doi.org/10.1093/nar/gkp272

Aukerman, M. J., & Sakai, H. (2003). Regulation of Flowering Time and Floral Organ Identity by a MicroRNA and Its APETALA2-Like Target Genes. *Plant Cell*, *15*(11), 2730–2741. https://doi.org/10.1105/tpc.016238

Beauclair, L., Yu, A., & Bouché, N. (2010). MicroRNA-directed cleavage and translational repression of the copper chaperone for superoxide dismutase mRNA in Arabidopsis. *Plant Journal*, *62*(3), 454–462. https://doi.org/10.1111/j.1365-313X.2010.04162.x

Bouché, N. (2010). New insights into miR398 functions in Arabidopsis. *Plant Signaling and Behavior*, *5*(6), 684–686. https://doi.org/10.4161/psb.5.6.11573

Brousse, C., Liu, Q., Beauclair, L., Deremetz, A., Axtell, M. J., & Bouché, N. (2014). A non-canonical plant microRNA target site. *Nucleic Acids Research*, *42*(8), 5270–5279. https://doi.org/10.1093/nar/gku157

Chen, X. (2004). A MicroRNA as a Translational Repressor of APETALA2 in Arabidopsis Flower Development. *Science*, *303*(5666), 2022–2025. https://doi.org/10.1126/science.1088060

Chen, Z. H., Bao, M. L., Sun, Y. Z., Yang, Y. J., Xu, X. H., Wang, J. H., Han, N., Bian, H. W., & Zhu, M. Y. (2011). Regulation of auxin response by miR393-targeted transport inhibitor response protein 1 is involved in normal development in Arabidopsis. *Plant Molecular Biology*, *77*(6), 619–629. https://doi.org/10.1007/s11103-011-9838-1

Chow, H. T., & Ng, D. W. K. (2017). Regulation of miR163 and its targets in defense against Pseudomonas syringae in Arabidopsis thaliana. *Scientific Reports*, *7*(January), 1–15. https://doi.org/10.1038/srep46433

Debernardi, J. M., Rodriguez, R. E., Mecchia, M. A., & Palatnik, J. F. (2012). Functional specialization of the plant miR396 regulatory network through distinct microRNA-target interactions. *PLoS Genetics*, *8*(1). https://doi.org/10.1371/journal.pgen.1002419

Dugas, D. V., & Bartel, B. (2008). Sucrose induction of Arabidopsis miR398 represses two Cu/Zn superoxide dismutases. *Plant Molecular Biology*, *67*(4), 403–417. https://doi.org/10.1007/s11103-008-9329-1

Emery, J. F., Floyd, S. K., Alvarez, J., Eshed, Y., Hawker, N. P., Izhaki, A., Baum, S. F., & Bowman, J. L. (2003). Radial Patterning of Arabidopsis Shoots by Class III HD-ZIP and KANADI Genes. *Current Biology*, *13*(20), 1768–1774. https://doi.org/10.1016/j.cub.2003.09.035

Fahlgren, N., Montgomery, T. A., Howell, M. D., Allen, E., Dvorak, S. K., Alexander, A. L., & Carrington, J. C. (2006). Regulation of AUXIN RESPONSE FACTOR3 by TAS3 ta-siRNA Affects Developmental Timing and Patterning in Arabidopsis. *Current Biology*, *16*(9), 939–944. https://doi.org/10.1016/j.cub.2006.03.065

Feng, Y. Z., Yu, Y., Zhou, Y. F., Yang, Y. W., Lei, M. Q., Lian, J. P., He, H., Zhang, Y. C., Huang, W., & Chen, Y. Q. (2020). A natural variant of mir397mediates a feedback loop in circadian rhythm. *Plant Physiology*, *182*(1), 204–214. https://doi.org/10.1104/pp.19.00710

Franco-Zorrilla, J. M., Valli, A., Todesco, M., Mateos, I., Puga, M. I., Rubio-Somoza, I., Leyva, A., Weigel, D., García, J. A., & Paz-Ares, J. (2007). Target mimicry provides a new mechanism for regulation of microRNA activity. *Nature Genetics*, *39*(8), 1033–1037. https://doi.org/10.1038/ng2079

Gao, R., Wang, Y., Gruber, M. Y., & Hannoufa, A. (2018). MiR156/SPL10 modulates lateral root development, branching and leaf morphology in Arabidopsis by silencing AGAMOUS-LIKE 79. *Frontiers in Plant Science*, *8*(January), 1–12. https://doi.org/10.3389/fpls.2017.02226

German, M. a, Pillay, M., Jeong, D.-H., Hetawal, A., Luo, S., Janardhanan, P., Kannan, V., Rymarquis, L. a, Nobuta, K., German, R., De Paoli, E., Lu, C., Schroth, G., Meyers, B. C., & Green, P. J. (2008). Global identification of microRNA-target RNA pairs by parallel analysis of RNA ends. *Nature Biotechnology*, *26*(8), 941–946. https://doi.org/10.1038/nbt1417

Guo, H.-S., Xie, Q., Fei, J.-F., & Chua, N.-H. (2005). MicroRNA directs mRNA cleavage of the transcription factor NAC1 to downregulate auxin signals for arabidopsis lateral root development. *The Plant Cell*, *17*(5), 1376–1386. https://doi.org/10.1105/tpc.105.030841

Harvey, J. J. W., Lewsey, M. G., Patel, K., Westwood, J., Heimstädt, S., Carr, J. P., & Baulcombe, D. C. (2011). An antiviral defense role of AGO2 in plants. *PLoS ONE*, *6*(1), 1–6. https://doi.org/10.1371/journal.pone.0014639

He, J., Xu, M., Willmann, M. R., McCormick, K., Hu, T., Yang, L., Starker, C. G., Voytas, D. F., Meyers, B. C., & Poethig, R. S. (2018). Threshold-dependent repression of SPL gene expression by miR156/miR157 controls vegetative phase change in Arabidopsis thaliana. *PLoS Genetics*, *14*(4), 1–28. https://doi.org/10.1371/journal.pgen.1007337

Hewezi, T., Piya, S., Qi, M., Balasubramaniam, M., Rice, J. H., & Baum, T. J. (2016). Arabidopsis miR827 mediates post-transcriptional gene silencing of its ubiquitin E3 ligase target gene in the syncytium of the cyst nematode Heterodera schachtii to enhance susceptibility. *Plant Journal*, *88*(2), 179–192. https://doi.org/10.1111/tpj.13238

Howell, M. D., Fahlgren, N., Chapman, E. J., Cumbie, J. S., Sullivan, C. M., Givan, S. A., Kasschau, K. D., & Carrington, J. C. (2007). Genome-wide analysis of the RNA-DEPENDENT RNA POLYMERASE6/DICER-LIKE4 pathway in Arabidopsis reveals dependency on miRNA- and tasiRNA-directed targeting. *Plant Cell*, *19*(3), 926–942. https://doi.org/10.1105/tpc.107.050062

Hunter, C., Willmann, M. R., Wu, G., Yoshikawa, M., Gutiérrez-Nava, M. de la L., & Poethig, R. S. (2006). Trans-acting siRNA-mediated repression of ETTIN and ARF4 regulates heteroblasty in Arabidopsis. *Development*, *133*(15), 2973–2981. https://doi.org/10.1242/dev.02491

Jones-Rhoades, M. W., & Bartel, D. P. (2004). Computational identification of plant MicroRNAs and their targets, including a stress-induced miRNA. *Molecular Cell*, *14*(6), 787–799. https://doi.org/10.1016/j.molcel.2004.05.027

Kasschau, K. D., Xie, Z., Allen, E., Llave, C., Chapman, E. J., Krizan, K. A., & Carrington, J. C. (2003). P1/HC-Pro, a viral suppressor of RNA silencing, interferes with Arabidopsis development and miRNA function. *Developmental Cell*, *4*(2), 205–217. https://doi.org/10.1016/S1534-5807(03)00025-X

Kawashima, C. G., Yoshimoto, N., Maruyama-Nakashita, A., Tsuchiya, Y. N., Saito, K., Takahashi, H., & Dalmay, T. (2009). Sulphur starvation induces the expression of microRNA-395 and one of its target genes but in different cell types. *Plant Journal*, *57*(2), 313–321. https://doi.org/10.1111/j.1365-313X.2008.03690.x

Kim, J. H., Woo, R. H., Kim, J., Lim, P. O., Lee, C. I., & Choi, S. H. (2009). Trifurcate Feed-Forward Regulation of Age-Dependent Cell Death Involving miR164 in Arabidopsis. *Science*, *323*(5917), 1053–1057. https://doi.org/DOI: 10.1126/science.1166386

Kutter, C., Schöb, H., Stadler, M., Meins, F., & Si-Ammour, A. (2007). MicroRNA-mediated regulation of stomatal development in Arabidopsis. *Plant Cell*, *19*(8), 2417–2429. https://doi.org/10.1105/tpc.107.050377

Lal, S., Pacis, L. B., & Smith, H. M. S. (2011). Regulation of the SQUAMOSA PROMOTER-BINDING PROTEIN-LIKE genes/microRNA156 Module by the homeodomain proteins PENNYWISE and POUND-FOOLISH in arabidopsis. *Molecular Plant*, *4*(6), 1123–1132. https://doi.org/10.1093/mp/ssr041

Lee, H., Yoo, S. J., Lee, J. H., Kim, W., Yoo, S. K., Fitzgerald, H., Carrington, J. C., & Ahn, J. H. (2010). Genetic framework for flowering-time regulation by ambient temperature-responsive miRNAs in Arabidopsis. *Nucleic Acids Research*, *38*(9), 3081–3093. https://doi.org/10.1093/nar/gkp1240

Liang, G., He, H., Li, Y., Wang, F., & Yu, D. (2014). Molecular mechanism of microRNA396 mediating pistil development in Arabidopsis. *Plant Physiology*, *164*(1), 249–258. https://doi.org/10.1104/pp.113.225144

Liang, G., Yang, F., & Yu, D. (2010). MicroRNA395 mediates regulation of sulfate accumulation and allocation in Arabidopsis thaliana. *Plant Journal*, *62*(6), 1046–1057. https://doi.org/10.1111/j.1365-313X.2010.04216.x

Lin, L. L., Wu, C. C., Huang, H. C., Chen, H. J., Hsieh, H. L., & Juan, H. F. (2013). Identification of microRNA 395a in 24-epibrassinolide-regulated root growth of Arabidopsis thaliana using microRNA arrays. *International Journal of Molecular Sciences*, *14*(7), 14270–14286. https://doi.org/10.3390/ijms140714270

Lin, W. Y., Lin, Y. Y., Chiang, S. F., Syu, C., Hsieh, L. C., & Chiou, T. J. (2018). Evolution of microRNA827 targeting in the plant kingdom. *New Phytologist*, *217*(4), 1712–1725. https://doi.org/10.1111/nph.14938

Liu, P. P., Montgomery, T. A., Fahlgren, N., Kasschau, K. D., Nonogaki, H., & Carrington, J. C. (2007). Repression of AUXIN RESPONSE FACTOR10 by microRNA160 is critical for seed germination and post-germination stages. *Plant Journal*, *52*(1), 133–146. https://doi.org/10.1111/j.1365-313X.2007.03218.x

Liu, D., Song, Y., Chen, Z., & Yu, D. (2009). Ectopic expression of miR396 suppresses GRF target gene expression and alters leaf growth in Arabidopsis. *Physiologia Plantarum*, *136*(2), 223–236. https://doi.org/10.1111/j.1399-3054.2009.01229.x

Llave, C., Xie, Z., Kasschau, K. D., & Carrington, J. C. (2002). Cleavage of Scarecrow-like mRNA targets directed by a class of Arabidopsis miRNA. *Science*, *297*(5589), 2053–2056. https://doi.org/10.1126/science.1076311

Ma, C., Burd, S., & Lers, A. (2015). MiR408 is involved in abiotic stress responses in Arabidopsis. *Plant Journal*, *84*(1), 169–187. https://doi.org/10.1111/tpj.12999

Mallory, A. C., Dugas, D. V, Bartel, D. P., & Bartel, B. (2004a). MicroRNA Regulation of NAC-Domain Targets Is Required for Proper Formation and Separation of Adjacent Embryonic, Vegetative, and Floral Organs. *Current Biology*, *14*(12), 1035–1046. https://doi.org/10.1016/j.cub.2004.06.022.

Mallory, A. C., Reinhart, B. J., Jones-Rhoades, M. W., Tang, G., Zamore, P. D., Barton, M. K., & Bartel, D. P. (2004b). MicroRNA control of PHABULOSA in leaf development: Importance of pairing to the microRNA 5′ region. *EMBO Journal*, *23*(16), 3356–3364. https://doi.org/10.1038/sj.emboj.7600340

Mallory, A. C., Bartel, D. P., & Bartel, B. (2005). MicroRNA-Directed Regulation of Arabidopsis AUXIN RESPONSE FACTOR17 Is Essential for Proper Development and Modulates Expression of Early Auxin Response Genes. *Development*, *17*(May), 1–16. https://doi.org/10.1105/tpc.105.031716.1

Marin, E., Jouannet, V., Herz, A., Lokerse, A. S., Weijers, D., Vaucheret, H., Nussaume, L., Crespi, M. D., & Maizel, A. (2010). mir390, Arabidopsis TAS3 tasiRNAs, and their AUXIN RESPONSE FACTOR targets define an autoregulatory network quantitatively regulating lateral root growth. *Plant Cell*, *22*(4), 1104–1117. https://doi.org/10.1105/tpc.109.072553

Mathieu, J., Yant, L. J., Murdter, F., Kuttner, F., & Schmid, M. (2009). Repression of Flowering by the miR172 Target SMZ. *PloS One*, *7*(7), e1000148. https://doi.org/10.1371/journal.pbio.1000148

Morea, E. G. O., da Silva, E. M., e Silva, G. F. F., Valente, G. T., Barrera Rojas, C. H., Vincentz, M., & Nogueira, F. T. S. (2016). Functional and evolutionary analyses of the miR156 and miR529 families in land plants. *BMC Plant Biology*, *16*(1), 1–13. https://doi.org/10.1186/s12870-016-0716-5

Ng, D. W. K., Zhang, C., Miller, M., Palmer, G., Whiteley, M., Tholl, D., & Jeffrey Chena, Z. (2011). Cis-and trans-regulation of miR163 and target genes confersnatural variation of secondary metabolites in two Arabidopsis species and their allopolyploids. *Plant Cell*, *23*(5), 1729–1740. https://doi.org/10.1105/tpc.111.083915

Ochando, I., Jover-Gil, S., Ripoll, J. J., Candela, H., Vera, A., Ponce, M. R., Martínez-Laborda, A., & Micol, J. L. (2006). Mutations in the microRNA complementarity site of the INCURVATA4 gene perturb meristem function and adaxialize lateral organs in Arabidopsis. *Plant Physiology*, *141*(2), 607–619. https://doi.org/10.1104/pp.106.077149

Palatnik, J. F., Wollmann, H., Schommer, C., Schwab, R., Boisbouvier, J., Rodriguez, R., Warthmann, N., Allen, E., Dezulian, T., Huson, D., Carrington, J. C., & Weigel, D. (2007). Sequence and Expression Differences Underlie Functional Specialization of Arabidopsis MicroRNAs miR159 and miR319. *Developmental Cell*, *13*(1), 115–125. https://doi.org/10.1016/j.devcel.2007.04.012

Peragine, A., Yoshikawa, M., Wu, G., Albrecht, H. L., & Poethig, R. S. (2004). SGS3 and SGS2/SDE1/RDR6 are required for juvenile development and the production of trans-acting siRNAs in Arabidopsis. *Genes and Development*, *18*(19), 2368–2379. https://doi.org/10.1101/gad.1231804

Rajagopalan, R., Vaucheret, H., Trejo, J., & Bartel, D. P. (2006). A diverse and evolutionarily fluid set of microRNAs in Arabidopsis thaliana. *Genes and Development*, *20*(24), 3407–3425. https://doi.org/10.1101/gad.1476406

Schwab, R., Palatnik, J. F., Riester, M., Schommer, C., Schmid, M., & Weigel, D. (2005). Specific effects of microRNAs on the plant transcriptome. *Developmental Cell*, *8*(4), 517–527. https://doi.org/10.1016/j.devcel.2005.01.018

Si-Ammour, A., Windels, D., Arn-Bouldoires, E., Kutter, C., Ailhas, J., Meins, F., & Vazquez, F. (2011). miR393 and secondary siRNAs regulate expression of the TIR1/AFB2 auxin receptor clade and auxin-related development of Arabidopsis leaves. *Plant Physiology*, *157*(2), 683–691. https://doi.org/10.1104/pp.111.180083

Song, J. B., Huang, S. Q., Dalmay, T., & Yang, Z. M. (2012). Regulation of leaf morphology by MicroRNA394 and its target LEAF CURLING RESPONSIVENESS. *Plant and Cell Physiology*, *53*(7), 1283–1294. https://doi.org/10.1093/pcp/pcs080

Sunkar, R., Kapoor, A., & Zhu, J. K. (2006). Erratum: Posttranscriptional induction of two Cu/Zn superoxide dismutase genes in Arabidopsis is mediated by downregulation of miR398 and important for oxidative stress tolerance (Plant Cell (2006) 18 (2051-2065)). *Plant Cell*, *18*(9), 2415. https://doi.org/10.1105/tpc.106.180960

Szaker, H. M., Darkó, É., Medzihradszky, A., Janda, T., Liu, H. C., Charng, Y. Y., & Csorba, T. (2019). miR824/AGAMOUS-LIKE16 Module Integrates Recurring Environmental Heat Stress Changes to Fine-Tune Poststress Development. *Frontiers in Plant Science*, *10*(November), 1–19. https://doi.org/10.3389/fpls.2019.01454

Tang, G., Reinhart, B. J., Bartel, D. P., & Zamore, P. D. (2003). A biochemical framework for RNA silencing in plants. *Genes and Development*, *17*(1), 49–63. https://doi.org/10.1101/gad.1048103

Thatcher, S. R., Burd, S., Wright, C., Lers, A., & Green, P. J. (2015). Differential expression of miRNAs and their target genes in senescing leaves and siliques: Insights from deep sequencing of small RNAs and cleaved target RNAs. *Plant, Cell and Environment*, *38*(1), 188–200. https://doi.org/10.1111/pce.12393

Vaucheret, H., Vazquez, F., Crété, P., & Bartel, D. P. (2004). The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. *Genes and Development*, *18*(10), 1187–1197. https://doi.org/10.1101/gad.1201404

Vazquez, F., Gasciolli, V., Crété, P., & Vaucheret, H. (2004). The Nuclear dsRNA Binding Protein HYL1 Is Required for MicroRNA Accumulation and Plant Development, but Not Posttranscriptional Transgene Silencing. *Current Biology*, *14*(4), 346–351. https://doi.org/10.1016/j.cub.2004.01.035

Vidal, E. A., Araus, V., Lu, C., Parry, G., Green, P. J., Coruzzi, G. M., & Gutiérrez, R. A. (2010). Nitrate-responsive miR393/AFB3 regulatory module controls root system architecture in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(9), 4477–4482. https://doi.org/10.1073/pnas.0909571107

Wang, J. W., Wang, L. J., Mao, Y. B., Cai, W. J., Xue, H. W., & Chen, X. Y. (2005). Control of root cap formation by MicroRNA-targeted auxin response factors in Arabidopsis. *Plant Cell*, *17*(8), 2204–2216. https://doi.org/10.1105/tpc.105.033076

Wang, L., Mai, Y. X., Zhang, Y. C., Luo, Q., & Yang, H. Q. (2010). MicroRNA171c-targeted SCL6-II, SCL6-III, and SCL6-IV genes regulate shoot branching in arabidopsis. *Molecular Plant*, *3*(5), 794–806. https://doi.org/10.1093/mp/ssq042

Wei, S., Gruber, M. Y., Yu, B., Gao, M. J., Khachatourians, G. G., Hegedus, D. D., Parkin, I. A. P., & Hannoufa, A. (2012). Arabidopsis mutant sk156 reveals complex regulation of SPL15 in a miR156-controlled gene network. *BMC Plant Biology*, *12*. https://doi.org/10.1186/1471-2229-12-169

Wójcik, A. M., Nodine, M. D., & Gaj, M. D. (2017). MiR160 and miR166/165 contribute to the LEC2-mediated auxin response involved in the somatic embryogenesis induction in arabidopsis. *Frontiers in Plant Science*, *8*(December), 1–17. https://doi.org/10.3389/fpls.2017.02024

Wu, G., and Poethig, R. S. (2006). Temporal regulation of shoot development in Arabidopsis thaliana by miR156 and its target SPL3. *Development*, *133*(18), 3539–3547. https://doi.org/10.1242/dev.02521

Wu, M. F., Tian, Q., & Reed, J. W. (2006). Arabidopsis microRNA 167 controls patterns of ARF6 and ARF8 expression, and regulates both female and male reproduction. *Development*, *133*(21), 4211–4218. https://doi.org/10.1242/dev.02602

Xie, Z., Allen, E., Fahlgren, N., Calamar, A., Givan, S. A., & Carrington, J. C. (2005). Expression of Arabidopsis MIRNA genes. *Plant Physiology*, *138*(4), 2145–2154. https://doi.org/10.1104/pp.105.062943

Xie, Z., Kasschau, K. D., & Carrington, J. C. (2003). Negative Feedback Regulation of Dicer-Like1 in Arabidopsis by microRNA-Guided mRNA Degradation. *Current Biology*, *13*(9), 784–789. https://doi.org/doi.org/10.1016/S0960-9822(03)00281-1

Xu, M. Y., Zhang, L., Li, W. W., Hu, X. L., Wang, M. B., Fan, Y. L., Zhang, C. Y., & Wang, L. (2014). Stress-induced early flowering is mediated by miR169 in Arabidopsis thaliana. *Journal of Experimental Botany*, *65*(1), 89–101. https://doi.org/10.1093/jxb/ert353

Xu, M., Hu, T., Zhao, J., Park, M. Y., Earley, K. W., Wu, G., Yang, L., & Poethig, R. S. (2016). Developmental Functions of miR156-Regulated SQUAMOSA PROMOTER BINDING PROTEIN-LIKE (SPL) Genes in Arabidopsis thaliana. *PLoS Genetics*, *12*(8), 1–29. https://doi.org/10.1371/journal.pgen.1006263

Yamasaki, H., Abdel-Ghany, S. E., Cohu, C. M., Kobayashi, Y., Shikanai, T., & Pilon, M. (2007). Regulation of copper homeostasis by micro-RNA in Arabidopsis. *Journal of Biological Chemistry*, *282*(22), 16369–16378. https://doi.org/10.1074/jbc.M700138200

Yan, J., Gu, Y., Jia, X., Kang, W., Pan, S., Tang, X., Chen, X., & Tang, G. (2012). Effective small RNA destruction by the expression of a short tandem target mimic in Arabidopsis. *The Plant Cell*, *24*(2), 415–427. https://doi.org/10.1105/tpc.111.094144

Yan, J., Gu, Y., Jia, X., Kang, W., Pan, S., Tang, X., Chen, X., & Tang, G. (2012). Effective small RNA destruction by the expression of a short tandem target mimic in Arabidopsis. *Plant Cell*, *24*(2), 415–427. https://doi.org/10.1105/tpc.111.094144

Yang, C. Y., Huang, Y. H., Lin, C. P., Lin, Y. Y., Hsu, H. C., Wang, C. N., Liu, L. Y. D., Shen, B. N., & Lin, S. S. (2015). MicroRNA396-targeted SHORT VEGETATIVE PHASE is required to repress flowering and is related to the development of abnormal flower symptoms by the phyllody symptoms1 effector1. *Plant Physiology*, *168*(4), 1702–1716. https://doi.org/10.1104/pp.15.00307

Yu, N., Niu, Q. W., Ng, K. H., & Chua, N. H. (2015). The role of miR156/SPLs modules in Arabidopsis lateral root development. *Plant Journal*, *83*(4), 673–685. https://doi.org/10.1111/tpj.12919

Zhang, H., He, H., Wang, X., Wang, X., Yang, X., Li, L., & Deng, X. W. (2011). Genome-wide mapping of the HY5-mediated genenetworks in Arabidopsis that involve both transcriptional and post-transcriptional regulation. *Plant Journal*, *65*(3), 346–358. https://doi.org/10.1111/j.1365-313X.2010.04426.x

Zhao, Y., Lin, S., Qiu, Z., Cao, D., Wen, J., Deng, X., Wang, X., Lin, J., & Li, X. (2015). MicroRNA857 is involved in the regulation of secondary growth of vascular tissues in Arabidopsis. *Plant Physiology*, *169*(4), 2539–2552. https://doi.org/10.1104/pp.15.01011

Zheng, Y., Li, Y. F., Sunkar, R., & Zhang, W. (2012). SeqTar: An effective method for identifying microRNA guided cleavage sites from degradome of polyadenylated transcripts in plants. *Nucleic Acids Research*, *40*(4). https://doi.org/10.1093/nar/gkr1092

References for Table S5

Boccara, M., Sarazin, A., Thiébeauld, O., Jay, F., Voinnet, O., Navarro, L., & Colot, V. (2014). The Arabidopsis miR472-RDR6 Silencing Pathway Modulates PAMP- and Effector-Triggered Immunity through the Post-transcriptional Control of Disease Resistance Genes. *PLOS Pathogens*, *10*(1), e1003883. https://doi.org/10.1371/journal.ppat.1003883

German, M. a, Pillay, M., Jeong, D.-H., Hetawal, A., Luo, S., Janardhanan, P., Kannan, V., Rymarquis, L. a, Nobuta, K., German, R., De Paoli, E., Lu, C., Schroth, G., Meyers, B. C., & Green, P. J. (2008). Global identification of microRNA-target RNA pairs by parallel analysis of RNA ends. *Nature Biotechnology*, *26*(8), 941–946. https://doi.org/10.1038/nbt1417

Park, Y. J., Lee, H. J., Kwak, K. J., Lee, K., Hong, S. W., & Kang, H. (2014). MicroRNA400-Guided Cleavage of Pentatricopeptide Repeat Protein mRNAs Renders Arabidopsis thaliana More Susceptible to Pathogenic Bacteria and Fungi. *Plant and Cell Physiology*, *55*(9), 1660–1668. https://doi.org/10.1093/pcp/pcu096

Sharma, D., Tiwari, M., Pandey, A., Bhatia, C., Sharma, A., & Trivedi, P. K. (2016). MicroRNA858 Is a Potential Regulator of Phenylpropanoid Pathway and Plant Development. *Plant Physiology*, *171*(2), 944–959. https://doi.org/10.1104/pp.15.01831

Table S1. Previously validated miRNA and tasiRNA targets found in *Arabidopsis thaliana*

*Brassicaceae and/*or *A. thaliana* specific miRNAs are highlighted in blue
miRNAs involved in stress response are highlighted in yellow
*miR163 and miR857 are both responsive to stress and *Brassicaceae* specific
**Analysis included the sequences for miR170/miR171a/miR171b
***The highest cleavage tag abundance found for this gene across all degradome libraries
****Identified as a HE target by TRUEE at a 20% library % cut-off
*****Expectation Score as determined by psRNATarget

				Library % Cut-off										
miRNA	**Gene ID**	**Target Description**	**10%**	**20%**	**30%**	**40%**	***Cat Score***	***Maximum Category***	**Cleavage tag abundance *****	**HE target at 20% cut-off******	**Expectation Score *******	**Validation method**	**Reference**	**sRNA sequence**
miR156	AT1G27360	SPL11, SQUAMOSA PROMOTER BINDING PROTEIN-LIKE 11	X	X	X	X	2.869	1	513.89	Yes	1	miRNA resistant target OE, miRNA KD, degradome	German et al., 2008; Zhang et al., 2011; Xu et al., 2016; He et al., 2018	UGACAGAAGAGAGUGAGCAC
miR156	AT1G27370	SPL10, SQUAMOSA PROMOTER BINDING PROTEIN-LIKE 10	X	X	X	X	2.295	1	267.76	Yes	1	miRNA resistant target OE, miRNA KD, degradome	Vazquez et al., 2004; German et al., 2008; Xu et al., 2016; Gao et al., 2018; He et al., 2018	UGACAGAAGAGAGUGAGCAC
miR156	AT1G53160	SPL4, SQUAMOSA PROMOTER BINDING PROTEIN-LIKE 4	X	X	X		1.51	1	97.17	Yes	1	5' RACE, miRNA resistant target OE, Degradome	Wu & Poethig, 2006; Lal et al., 2011	UGACAGAAGAGAGUGAGCAC
miR156	AT1G69170	SPL6,SQUAMOSA PROMOTER BINDING PROTEIN-LIKE 6	X	X	X	X	2.016	1	56.01	Yes	1	correlation of miRNA/target mRNA levels, miRNA OE, degradome	German et al., 2008; Xu et al., 2016; He et al., 2018	UGACAGAAGAGAGUGAGCAC
miR156	AT2G33810	SPL3, SQUAMOSA PROMOTER BINDING PROTEIN-LIKE 3	X	X	X	X	3.00	1	213.52	Yes	0.5	5' RACE, miRNA resistant target OE, miRNA KD, degradome	German et al., 2008; Wu & Poethig, 2006; Xu et al., 2016; He et al., 2018	UGACAGAAGAGAGUGAGCAC
miR156	AT2G42200	SPL9,SQUAMOSA PROMOTER BINDING PROTEIN-LIKE 9	X	X	X		0.328	2	63.95	Yes	1	miRNA resistant target OE, miRNA KD, degradome	Addo-Quaye et al., 2008; Yu et al., 2015; Xu et al., 2016; He et al., 2018	UGACAGAAGAGAGUGAGCAC
miR156	AT3G15270	SPL5, SQUAMOSA PROMOTER BINDING PROTEIN-LIKE 5	X	X	X	X	1.852	1	72.59	Yes	2	5' RACE, miRNA resistant target OE, degradome	Wu & Poethig, 2006; German et al., 2008; Lee et al., 2010; Lal et al., 2011	UGACAGAAGAGAGUGAGCAC
miR156	AT3G57920	SPL15, SQUAMOSA PROMOTER-BINDING PROTEIN LIKE 15	X	X	X	X	1.443	1	59.19	Yes	1	miRNA resistant target OE, miRNA KD, degradome	Addo-Quaye et al., 2008; Wei et al., 2012; Morea et al., 2016; He et al., 2018	UGACAGAAGAGAGUGAGCAC

miR156	AT5G43270	SPL2, SQUAMOSA PROMOTER BINDING PROTEIN-LIKE 2	X	X	X	X	0.967	1	662.53	Yes	1	miRNA resistant target OE, miRNA KD, degradome	German et al., 2008; Lee et al., 2010; Xu et al., 2016; He et al., 2018	UGACAGAAGAGAGUGAGCAC
miR156	AT5G50570	SPL13, SQUAMOSA PROMOTER-BINDING PROTEIN LIKE 13	X	X	X	X	3.18	1	552.36	Yes	1	miRNA resistant target OE, miRNA KD, degradome	German et al., 2008; Xu et al., 2016; He et al., 2018	UGACAGAAGAGAGUGAGCAC
miR159	AT2G34010	TCP INTERACTOR CONTAINING EAR MOTIF PROTEIN 4, TIE4/MRG1	X	X	X		1.656	1	195.87	Yes	2.5	5' RACE, degradome	German et al., 2008; Alves-Junior et al., 2009	UUUGGAUUGAAGGGAGCUCUA
miR159	AT3G11440	MYB65, MYB Domain Protein 65	X	X	X	X	3.279	1	429.27	Yes	2.5	miRNA OE, miRNA resistant target, miRNA and target KO, degradome	Allen et al., 2007; Palatnik et al., 2007; German et al., 2008	UUUGGAUUGAAGGGAGCUCUA
miR159	AT4G27330	NOZZLE, NZZ, SPL, SPOROCYTELESS	NA	NA	NA	NA	NA	NA	NA	No	2.5	5' RACE, miRNA OE	Chorostecki et al., 2012	UUUGGAUUGAAGGGAGCUCUA
miR159	AT5G06100	MYB33, MYB Domain Protein 33	X	X	X	X	2.213	1	361.6	Yes	2.5	miRNA OE, miRNA resistant target, miRNA and target KO, degradome	Allen et al., 2007; Palatnik et al., 2007; German et al., 2008	UUUGGAUUGAAGGGAGCUCUA
miR159	AT5G55930	ATOPT1, OLIGOPEPTIDE TRANSPORTER 1, OPT1	NA	NA	NA	NA	NA	NA	NA	No	3.5	5' RACE, miRNA OE, miRNA KO	Schwab et al., 2005; Alonso-Peral et al., 2010	UUUGGAUUGAAGGGAGCUCUA
miR160	AT1G77850	ARF17, AUXIN RESPONSE FACTOR 17	X	X	X	X	4.148	1	839.3	Yes	0.5	5' RACE, miRNA resistant target OE, miRNA OE, Degradome	Wang et al., 2005; Mallory et al., 2005; German et al., 2008; Wójcik et al., 2017	UGCCUGGCUCCCUGUAUGCCA
miR160	AT2G28350	ARF10, AUXIN RESPONSE FACTOR 10	X	X	X	X	3.557	1	1489.44	Yes	1	5' RACE, miRNA resistant target OE, miRNA OE, Degradome	Mallory et al., 2005; Wang et al., 2005; Liu et al., 2007; German et al., 2008	UGCCUGGCUCCCUGUAUGCCA
miR160	AT4G30080	ARF16, AUXIN RESPONSE FACTOR 16	X	X	X	X	4.115	1	725.3	Yes	1.5	5' RACE, miRNA resistant target OE, miRNA OE, Degradome	Wang et al., 2005; Mallory et al., 2005; German et al., 2008; Wójcik et al., 2017	UGCCUGGCUCCCUGUAUGCCA
miR161	AT1G62590	PPR-AC, Pentatricopeptide Adenylate Cyclase	NA	NA	NA	NA	NA	NA	NA	No	4.5	5' RACE, degradome	Howell et al., 2007; German et al., 2008	UGAAAGUGACUACAUCGGGGU

miR161	AT1G62910	RPF4, RNA Processing Factor 4	X	X	X	X	0.541	1	140.94	Yes	2	5' RACE, degradome	Howell et al., 2007; German et al., 2008	UGAAAGUGACUCACAUCGGGU
miR161	AT1G63080	PPR1, Pentatricopeptide Repeat 1	NA	NA	NA	NA	NA	NA	NA	No	2	5' RACE, degradome	Howell et al., 2007; German et al., 2008	UGAAAGUGACUCACAUCGGGU
miR161	AT1G63130	RPF6, RNA Processing Factor 6	X	X	X	X	0.541	1	140.94	Yes	2	5' RACE, degradome	Howell et al., 2007; German et al., 2008	UGAAAGUGACUCACAUCGGGU
miR161	AT1G63150	Tetratricopeptide Repeat (TPR)-like Superfamily Protein	NA	NA	NA	NA	NA	NA	NA	No	3	5' RACE, degradome	Allen et al., 2004; Howell et al., 2007; German et al., 2008	UGAAAGUGACUCACAUCGGGU
miR161	AT1G63400	Pentatricopeptide Repeat (PPR) Superfamily Protein	NA	NA	NA	NA	NA	NA	NA	No	2	5' RACE, degradome	Howell et al., 2007; German et al., 2008	UGAAAGUGACUCACAUCGGGU
miR162	AT1G01040	DCL1, DICER-LIKE1	X	X	X	X	1.36	1	116.84	Yes	3	5' RACE, degradome	Xie et al., 2003; Jones-Rhoades et al., 2004; German et al., 2008	UCGAUAAACCUCUGCAUCCAG
miR163*	AT1G15125	SAMT, S-adenosyl-L-methionine-dependent Methyltransferases Superfamily Protein	X				0.61	1	57.3	No	1	5' RACE, degradome	Xie et al., 2005; Addo-Quaye et al., 2008	UUGAAGAGGACUUGGAACUUCGAU
miR163*	AT1G66700	PXMT1, S-adenosyl-L-methionine-dependent Methyltransferases Superfamily Protein	X	X	X		1.54	1	70.6	Yes	2	5' RACE, miRNA KO, Deg	Xie et al., 2005; German et al., 2008; Ng et al., 2011; Chow et al., 2017	UUGAAGAGGACUUGGAACUUCGAU
miR163*	AT3G44860	FAMT, Farnesoic Acid Carboxyl-O-methyltransferase	X	X	X		1.49	1	320.44	Yes	3	5' RACE, miRNA KO	Allen et al., 2004; Xie et al., 2005; Ng et al., 2011; Chow et al., 2017	UUGAAGAGGACUUGGAACUUCGAU
miR164	AT1G56010	NAC1, NAC Domain Containing Protein 1	X	X	X	X	3.967	1	1955.33	Yes	1	5' RACE, degradome, miRNA resistant target OE, miRNA KO	Mallory et al., 2004a ; Guo et al., 2005; German et al., 2008	UGGAGAAGCAGGGCACGUGCA
miR164	AT3G15170	CUC1, CUP-SHAPED COTYLEDON 1	X				0.66	1	78.92	No	1	5' RACE, miRNA resistant target OE, miRNA OE, degradome	Kasschau et al., 2003; Mallory et al., 2004a; German et al., 2008	UGGAGAAGCAGGGCACGUGCA
miR164	AT5G07680	NAC80, NAC Domain Containing Protein 80, NAC4	X	X			0.74	1	20.67	Yes	0.5	5' RACE, degradome	Mallory et al., 2004a; German et al., 2008	UGGAGAAGCAGGGCACGUGCA

miR164	AT5G39610	NAC6, NAC Domain Containing Protein 6, NAC92	X	X	X	X	3.197	1	260.79	Yes	2	miRNA OE, miRNA resistant target, Degradome	German et al., 2008; Kim et al., 2009	UGGAGAAGCAGGGCACGUGCA
miR164	AT5G53950	CUC2, CUP-SHAPED COTYLEDON 2	X	X			0.984	1	213.52	Yes	1	5' RACE, miRNA OE, degradome	Kasschau et al., 2003; Mallory et al., 2004a; German et al., 2008	UGGAGAAGCAGGGCACGUGCA
miR164	AT5G61430	NAC100, NAC Domain Containing Protein 100	X	X	X		1.18	1	33.82	Yes	0.5	5' RACE, degradome	Mallory et al., 2004a; German et al., 2008	UGGAGAAGCAGGGCACGUGCA
miR166	AT1G30490	PHV, PHAVOLUTA	X	X	X	X	4.082	1	1140.94	Yes	1	miRNA resistant target and miRNA coexpression, 5' RACE, degradome, miRNA decoy OE	Tang et al., 2003; Mallory et al., 2004b; German et al., 2008; Yan et al., 2012	UCGGACCAGGCUUCAUUCCCC
miR166	AT1G52150	ICU4, INCURVATA 4	X	X	X	X	3.016	1	424.21	Yes	0.5	5' RACE, degradome, miRNA resistant target, miRNA decoy OE	Mallory et al., 2004b; Ochando et al., 2006; German et al., 2008; Yan et al., 2012	UCGGACCAGGCUUCAUUCCCC
miR166	AT2G34710	PHB, PHABULOSA	X	X	X	X	4.082	1	1223.06	Yes	0.5	miRNA resistant target, 5' RACE, degradome	Tang et al., 2003; Mallory et al., 2004b; German et al., 2008	UCGGACCAGGCUUCAUUCCCC
miR166	AT4G32880	HB-8, HOMEOBOX GENE 8	X	X	X	X	2.279	1	131.44	Yes	0.5	5' RACE, degradome	Mallory et al., 2004b; German et al., 2008	UCGGACCAGGCUUCAUUCCCC
miR166	AT5G60690	IFL1, INTERFASCICULAR FIBERLESS 1, REV, REVOLUTA	X	X	X	X	3.311	1	295.07	Yes	0.5	5' RACE, degradome, miRNA resistant target, miRNA decoy OE	Emery et al., 2003; Mallory et al., 2004b; German et al., 2008; Yan et al., 2012	UCGGACCAGGCUUCAUUCCCC
miR167	AT1G30330	ARF6, AUXIN RESPONSE FACTOR 6	X	X	X	X	1.754	1	221.53	Yes	3.5	5' RACE, miRNA OE, miRNA resistant target OE, degradome	Jones-Rhoades et al., 2004; Wu et al., 2006; German et al., 2008	UGAAGCUGCCAGCAUGAUCUA
miR167	AT5G37020	ARF8, AUXIN RESPONSE FACTOR 8	X	X	X	X	1.279	1	226.42	Yes	3.5	5' RACE, miRNA OE, miRNA resistant target OE, degradome	Tang et al., 2003; Mallory et al., 2004b; German et al., 2008	UGAAGCUGCCAGCAUGAUCUA
miR168	AT1G48410	AGO1, ARGONAUTE 1	X	X	X	X	2.754	1	232.11	Yes	3	miRNA resistant target OE, Deg	Vaucheret et al., 2004; German et al., 2008	UCGCUUGGUGCAGGUCGGGAA
miR169	AT1G17590	NF-YA8, NUCLEAR FACTOR Y, SUBUNIT A8	X	X	X	X	3.344	1	201.11	Yes	1	miRNA OE, degradome	German et al., 2008; Xu et al., 2014	CAGCCAAGGAUGACUUGCCGA

miR169	AT1G54160	NF-YA5, NUCLEAR FACTOR Y, SUBUNIT A5	X	X	X	X	0.74	1	57.78	Yes	2	miRNA OE, degradome	German et al., 2008; Xu et al., 2014	CAGCCAAGGAUGACUUGCCGA
miR169	AT1G72830	NF-YA3, NUCLEAR FACTOR Y, SUBUNIT A3	X	X	X	X	3.049	1	139.35	Yes	2	miRNA OE, degradome	German et al., 2008; Xu et al., 2014	CAGCCAAGGAUGACUUGCCGA
miR169	AT3G05690	NF-YA2, NUCLEAR FACTOR Y, SUBUNIT A2	X	X	X	X	2.803	1	199.82	Yes	2.5	miRNA OE, miRNA resistant target OE, 5' RACE, degradome	Jones-Rhoades et al., 2004; German et al., 2008; Xu et al., 2014	CAGCCAAGGAUGACUUGCCGA
miR169	AT3G20910	NF-YA9, NUCLEAR FACTOR Y, SUBUNIT A9	X	X	X		1.18	1	68.58	Yes	2.5	miRNA OE, degradome	German et al., 2008; Xu et al., 2014	CAGCCAAGGAUGACUUGCCGA
miR169	AT5G06510	NF-YA10, NUCLEAR FACTOR Y, SUBUNIT A10	X	X	X	X	2.72	1	97.27	Yes	2.5	miRNA OE, degradome	German et al., 2008; Xu et al., 2014	CAGCCAAGGAUGACUUGCCGA
miR169	AT5G12840	NF-YA1, NUCLEAR FACTOR Y, SUBUNIT A1	X	X	X		1.30	1	23.66	Yes	2.5	miRNA OE, degradome	German et al., 2008; Xu et al., 2014	CAGCCAAGGAUGACUUGCCGA
miR170 / miR171 **	AT2G45160	HAM1, HAIRY MERISTEM 1	X	X	X	X	3.361	1	9581.42	Yes	0	degradome, miRNA OE, miRNA resistant target OE	German et al., 2008; Wang et al., 2010	UGAUUGAGCCGUGUCAAUAUC/UGAUUGAGCCGCGCCAAUAUC/UUGAGCCGUGCCAAUAUCACG
miR170 / miR171 **	AT3G60630	HAM2, HAIRY MERISTEM 2	X	X	X	X	1.41	1	1915.12	Yes	0	degradome, 5' RACE, miRNA OE, miRNA resistant target OE	Llave et al., 2002; German et al., 2008	UGAUUGAGCCGUGUCAAUAUC/UGAUUGAGCCGCGCCAAUAUC/UUGAGCCGUGCCAAUAUCACG
miR170 / miR171 **	AT4G00150	HAM3, HAIRY MERISTEM 3	X	X	X	X	1.607	1	276.2	Yes	0	degradome, 5' RACE, miRNA OE, miRNA resistant target OE	Llave et al., 2002; German et al., 2008	UGAUUGAGCCGUGUCAAUAUC/UGAUUGAGCCGCGCCAAUAUC/UUGAGCCGUGCCAAUAUCACG

miR172	AT2G28550	TOE1, TARGET OF EAT 1	X	X	X	X	2.902	1	461.2	Yes	0.5	5' RACE, miRNA OE, degradome	Aukerman & Sakai, 2003; Kasschau et al., 2003; German et al., 2008	AGAAUCUUG AUGAUGCUG CAU
miR172	AT2G39250	SNZ, SCHNARCHZAPFEN	X	X	X		1.13	1	111.44	Yes	0.5	5' RACE, degradome	German et al., 2008; Mathieu et al., 2009	AGAAUCUUG AUGAUGCUG CAU
miR172	AT3G54990	SMZ, SCHLAFMUTZE	X	X	X		1.38	1	51.71	Yes	1.5	5' RACE, miRNA resistant target OE, correlation of miRNA/target mRNA levels, degradome	German et al., 2008; Mathieu et al., 2009	AGAAUCUUG AUGAUGCUG CAU
miR172	AT4G36920	AP2, APETELA 2	NA	NA	NA	NA	NA	NA	NA	No	0.5	5' RACE, miRNA OE, hen1 and dcl1 mutant, degradome	Aukerman & Sakai, 2003; Kasschau et al., 2003; Chen, 2004; German et al., 2008	AGAAUCUUG AUGAUGCUG CAU
miR172	AT5G60120	TOE2, TARGET OF EAT 2	X	X	X	X	4.016	1	1040.46	Yes	0.5	5' RACE, miRNA OE, correlation of miRNA/target mRNA levels, degradome	Kasschau et al., 2003; German et al., 2008; Lee et al., 2010	AGAAUCUUG AUGAUGCUG CAU
miR172	AT5G67180	TOE3, TARGET OF EAT 3	X	X	X	X	2.656	1	207.63	Yes	0.5	5' RACE, degradome	German et al., 2008	AGAAUCUUG AUGAUGCUG CAU
miR319	AT1G30210	TCP24, TEOSINTE BRANCHED 24, cycloidea and PCF transcription factor 24	X	X	X	X	1.639	1	215.15	Yes	2.5	5' RACE, miRNA OE, degradome	Palatnik et al., 2003; German et al., 2008	UUGGACUGA AGGGAGCUC CCU
miR319	AT1G53230	TCP3, TEOSINTE BRANCHED 1, cycloidea and PCF transcription factor 3	X	X	X	X	1.23	1	167.99	Yes	3	5' RACE, miRNA OE, degradome	Palatnik et al., 2003; German et al., 2008	UUGGACUGA AGGGAGCUC CCU
miR319	AT2G31070	TCP10, TEOSINTE BRANCHED 10, TCP domain protein 10	X	X			0.28	1	106.76	Yes	2.5	5' RACE, miRNA OE, degradome	Palatnik et al., 2003; German et al., 2008	UUGGACUGA AGGGAGCUC CCU
miR319	AT3G15030	TCP4, TEOSINTE BRANCHED 4, TCP family transcription factor 4	X	X	X	X	2.23	1	126.83	Yes	2.5	5' RACE, miRNA OE, miRNA resistant target OE, degradome	Palatnik et al., 2003; German et al., 2008	UUGGACUGA AGGGAGCUC CCU

miR319	AT4G18390	TCP2, TEOSINTE BRANCHED 1, cycloidea and PCF transcription factor 2	X	X	X		0.951	1	253.67	Yes	2.5	5' RACE, miRNA OE, degradome	Palatnik et al., 2003; German et al., 2008	UUGGACUGAAGGGAGCUCCCU
miR393	AT1G12820	AFB3, AUXIN SIGNALING F BOX PROTEIN 3	X	X	X	X	4.016	1	731.71	Yes	1	5' RACE, correlation of miRNA/target mRNA levels, miRNA KO, degradome, miRNA OE	Jones-Rhoades et al., 2004; German et al., 2008; Vidal et al., 2009; Chen et al., 2011; Si-Ammour et al., 2011	UCCAAAGGGAUCGCAUUGAUCC
miR393	AT3G23690	CIB1 LIKE PROTEIN 2, bHLH	X	X	X		0.69	1	24.69	Yes	2.5	5' RACE, degradome	Jones-Rhoades et al., 2004; German et al., 2008	UCCAAAGGGAUCGCAUUGAUCC
miR393	AT3G26810	AFB2, AUXIN SIGNALING F BOX PROTEIN 2	X	X	X	X	3.066	1	227.4	Yes	1	5' RACE, miRNA KO, miRNA OE, degradome	Jones-Rhoades et al., 2004; German et al., 2008; Chen et al., 2011; Si-Ammour et al., 2011	UCCAAAGGGAUCGCAUUGAUCC
miR393	AT3G62980	TIR1, TRANSPORT INHIBITOR RESPONSE 1	X	X	X	X	3.066	1	247.69	Yes	1	5' RACE, miRNA OE, miRNA resistant target, degradome	Jones-Rhoades et al., 2004; German et al., 2008; Chen et al., 2011	UCCAAAGGGAUCGCAUUGAUCC
miR393	AT4G03190	AFB1, AUXIN SIGNALING F BOX PROTEIN 1	X	X	X	X	3.328	1	228.53	Yes	2	5' RACE, degradome	Jones-Rhoades et al., 2004; German et al., 2008	UCCAAAGGGAUCGCAUUGAUCC
miR394	AT1G27340	LCR, LEAF CURLING RESPONSIVENESS	X	X	X		0.48	1	33.16	Yes	1	5' RACE, miRNA OE, miRNA decoy OE, miRNA resistant target, degradome	Jones-Rhoades et al., 2004; German et al., 2008; Song et al., 2012	UUGGCAUUCUGUCCACCUCC
miR395	AT3G22890	APS1, ATP SULFURYLASE 1, ATPS1	X	X	X	X	1.787	1	350.01	Yes	3	5' RACE, miRNA OE, degradome	German et al., 2008; Kawashima et al., 2009; Liang et al., 2010	CUGAAGUGUUUGGGGGAACUC
miR395	AT4G14680	APS3, ATP-SULFURYLASE 3, ATPS3	NA	NA	NA	NA	NA	NA	NA	No	3	5' RACE, miRNA OE	Kawashima et al., 2009; Liang et al., 2010	CUGAAGUGUUUGGGGGAACUC
miR395	AT5G10180	SULTR2;1, SULFATE TRANSPORTER 2;1	X				0.20	1	20.78	No	1.5	5' RACE, miRNA OE, miRNA loss of function	Kawashima et al., 2009; Liang et al., 2010	CUGAAGUGUUUGGGGGAACUC
miR395	AT5G13630	GUN5, GENOMES UNCOUPLED 5	X	X	X	X	0.51	2	72.7	Yes	2	miRNA KO, miRNA-target co-expression	Lin et al., 2013	CUGAAGUGUUUGGGGGAACUC
miR395	AT5G43780	APS4, ATP-SULFURYLASE 4, ATPS4	X	X			0.59	1	39.63	Yes	1.5	5' RACE, miRNA OE	Jones-Rhoades et al., 2004; Kawashima et al., 2009; Liang et al., 2010	CUGAAGUGUUUGGGGGAACUC

miR396	AT1G10120	CIB4, CRY2-INTERACTING BHLH74	X	X	X	X	2.656	1	327.74	Yes	2.5	5' RACE, miRNA OE, ago1 mutant, degradome	German et al., 2008; Debernardi et al., 2012	UUCCACAGCUUUCUUGAACUG
miR396	AT1G53910	RAP2.12, RELATED TO AP2 12	X	X	X	X	2.951	1	226.42	Yes	3	5' RACE, degradome	Zheng et al., 2012	UUCCACAGCUUUCUUGAACUG
miR396	AT2G22840	GRF1, GROWTH-REGULATING FACTOR 1	X	X	X	X	4.016	1	1727.66	Yes	3	5' RACE, miRNA decoy OE, miRNA OE	Jones-Rhoades et al., 2004; German et al., 2008; Liu et al., 2009; Liang et al., 2014	UUCCACAGCUUUCUUGAACUG
miR396	AT2G36400	GRF3, GROWTH-REGULATING FACTOR 3	X	X	X	X	3.279	1	443.16	Yes	3	5' RACE, miRNA decoy OE, miRNA OE	Jones-Rhoades et al., 2004; German et al., 2008; Liu et al., 2009; Liang et al., 2014	UUCCACAGCUUUCUUGAACUG
miR396	AT2G45480	GRF9, GROWTH-REGULATING FACTOR 9	X	X	X	X	1.89	1	128.71	Yes	3	5' RACE, miRNA decoy OE, miRNA OE, miRNA resistant target	Jones-Rhoades et al., 2004; German et al., 2008; Liang et al., 2014	UUCCACAGCUUUCUUGAACUG
miR396	AT3G14110	FLU, FLUORESCENT IN BLUE LIGHT	X	X	X	X	1.984	1	219.95	Yes	3	5' RACE, miRNA OE	Chorostecki et al., 2012	UUCCACAGCUUUCUUGAACUG
miR396	AT3G52910	GRF4, GROWTH-REGULATING FACTOR 4	X	X	X	X	1.49	1	34.76	Yes	3	5' RACE, miRNA OE	Liang et al., 2014	UUCCACAGCUUUCUUGAACUG
miR396	AT4G24150	GRF8, GROWTH-REGULATING FACTOR 8	X	X	X	X	2.656	1	1086.06	Yes	3	5' RACE, miRNA decoy OE, miRNA OE	Jones-Rhoades et al., 2004; German et al., 2008; Liang et al., 2014	UUCCACAGCUUUCUUGAACUG
miR396	AT4G37740	GRF2, GROWTH-REGULATING FACTOR 2	X	X	X	X	4	1	1216.29	Yes	3	5' RACE, miRNA decoy OE, miRNA OE	Jones-Rhoades et al., 2004; German et al., 2008; Liu et al., 2009; Liang et al., 2014	UUCCACAGCUUUCUUGAACUG
miR396	AT5G43060	MMG4.7, RD21B, RESPONSIVE TO DEHYDRATION 21B	X	X	X		0.67	1	46.02	Yes	2	5' RACE, miRNA OE, degradome	Chorostecki et al., 2012; Yang et al., 2015	UUCCACAGCUUUCUUGAACUG
miR396	AT5G53660	GRF7, GROWTH-REGULATING FACTOR 7	X				0.12	2	34.62	No	3	5' RACE, miRNA OE miRNA resistant target	Jones-Rhoades et al., 2004; Liang et al., 2014	UUCCACAGCUUUCUUGAACUG
miR397	AT2G29130	LAC2, LACCASE 2	X	X			1.12	1	42.06	Yes	0.5	5' RACE, correlation of miRNA/target mRNA levels, degradome	Jones-Rhoades et al., 2004; Abdel-Ghany & Pilon, 2008; German et al., 2008	UCAUUGAGUGCAGCGUUGAUG

miR397	AT3G60250	CKB3, CASEIN KINASE II BETA CHAIN 3	X	X	X	X	1.00	1	192.6	Yes	3	5' RACE, miRNA OE, degradome	German et al., 2008; Feng et al., 2020	UCAUUGAGUGCAGCGUUGAUG
miR397	AT5G60020	LAC17, LACCASE 17	NA	NA	NA	NA	NA	NA	NA	No	1	5' RACE, correlation of miRNA/target mRNA levels, degradome	Jones-Rhoades et al., 2004; Abdel-Ghany & Pilon, 2008; German et al., 2008	UCAUUGAGUGCAGCGUUGAUG
miR398	AT1G08830	CSD1 COPPER/ZINC SUPEROXIDE DISMUTASE 1, SOD1	X	X	X	X	3.279	1	724.36	Yes	4	5' RACE, correlation of miRNA/target mRNA levels, miRNA resistant target, miRNA OE	Jones-Rhoades et al., 2004; Sunkar et al., 2006; Dugas & Bartel., 2008	UGUGUUCUCAGGUCACCCCUU
miR398	AT1G12520	CCS1, COPPER CHAPERONE FOR SOD1	X	X	X	X	3.574	1	2490.65	Yes	3.5	5' RACE, miRNA resistant target, miRNA OE, miRNA KO	Beauclair et al., 2010; Bouché, 2010	UGUGUUCUCAGGUCACCCCUU
miR398	AT2G28190	CSD2 COPPER/ZINC SUPEROXIDE DISMUTASE 2, SOD2	X	X	X	X	2.672	1	674.57	Yes	5	5' RACE, correlation of miRNA/target mRNA levels, miRNA resistant target, miRNA OE	Jones-Rhoades et al., 2004; Sunkar et al., 2006; Dugas & Bartel., 2008	UGUGUUCUCAGGUCACCCCUU
miR398	AT3G15640	COX5b–1, cytochrome c oxidase, Rubredoxin-like superfamily protein	X	X	X	X	1.344	1	185.78	Yes	4	5' RACE, correlation of miRNA/target mRNA levels	Jones-Rhoades et al., 2004; Yamasaki et al., 2007	UGUGUUCUCAGGUCACCCCUU
miR398	AT3G27200	Cupredoxin superfamily protein	X	X	X	X	3.115	1	453.78	Yes	4	5' RACE, Degradome	Zheng et al., 2012; Brousse et al., 2014	UGUGUUCUCAGGUCACCCCUU
miR399	AT2G33770	PHO2, PHOSPHATE 2, UBC24, UBIQUITIN-CONJUGATING ENZYME 24	X	X			0.21	2	36.15	Yes	0	5' RACE, miRNA OE, decoy OE, degradome	Allen et al., 2005; Franco-Zorrilla et al. 2007; German et al., 2008	UGCCAAAGGAGAUUUGCCCUG
miR403	AT1G31280	AGO2, ARGONAUTE 2	X	X	X	X	3.918	1	636.99	Yes	0	5' RACE, loss of ago1 function, degradome	Allen et al., 2005; German et al., 2008; Harvey et al., 2011	UUAGAUUCACGCACAAACUCG
miR408	AT1G72230	CUPREDOXIN, Cupredoxin superfamily protein	X	X			0.754	1	475.39	Yes	4.5	miRNA OE, miRNA KO, degradome	German et al., 2008; Ma et al., 2015; Thatcher et al., 2015	AUGCACUGCCUCUUCCCUGGC
miR408	AT2G02850	ARPN, PLANTACYANIN	X	X	X	X	2.328	1	546.93	Yes	1	5' RACE, miRNA OE, miRNA KO, degradome	Abdel-Ghany & Pilon, 2008; German et al., 2008; Ma et al., 2015	AUGCACUGCCUCUUCCCUGGC

miR408	AT2G30210	LAC3, LACCASE 3	X				0.20	2	49.87	No	3.5	5' RACE, miRNA OE, miRNA KO, degradome	Abdel-Ghany & Pilon, 2008; Ma et al., 2015	AUGCACUGCCUCUUCCCUGGC
miR408	AT2G44790	UCC2, UCLACYANIN 2	X	X	X	X	1.607	1	806.61	Yes	4.5	miRNA OE, miRNA KO, degradome	German et al., 2008; Ma et al., 2015; Thatcher et al., 2015	AUGCACUGCCUCUUCCCUGGC
miR823	AT1G69770	CMT3, CHROMOMETHYLASE 3	X	X	X	X	2.25	1	106.42	Yes	1.5	5' RACE, degradome	Rajagopalan et al., 2006; German et al., 2008	UGGGUGGUGAUCAUAUAAGAU
miR824	AT3G57230	AGL16, AGAMOUS-LIKE 16	X	X	X	X	3.852	1	858.71	Yes	0.5	5' RACE, miRNA decoy OE, miRNA resistant target, degradome	Rajagopalan et al., 2006; Kutter et al., 2007; German et al., 2008; Szaker et al., 2019	UAGACCAUUUGUGAGAAGGGA
miR827	AT1G02860	NLA, NITROGEN LIMITATION ADAPTATION, SYG1	X				0.66	1	39.39	No	0	miRNA resistant target, target KO, Degradome	Addo-Quaye et al., 2008; Hewezi et al., 2016; Lin et al., 2018	UUAGAUGACCAUCAACAAACU
miR857*	AT3G09220	LAC7, LACCASE 7	X				0.25	1	10.39	No	0.5	5' RACE, correlation of miRNA/target mRNA levels, miRNA OE	Abdel-Ghany & Pilon, 2008; Zhao et al., 2015	UUUUGUAUGUUGAAGGUGUAU
TasiARFs	AT2G33860	ARF3, AUXIN RESPONSE FACTOR 3	X	X	X	X	1.836	1	169.48	Yes	0.5	5' RACE, *ago7, rdr6, sgs3, tas3* mutant, TasiRNA resistant target, degradome	Peragine et al., 2004; Allen et al., 2005; Fahlgren et al., 2006; Hunter et al., 2006; German et al., 2008; Marin et al., 2010	UUCUUGACCUUGUAAGACCUU
TasiARFs	AT5G60450	ARF4, AUXIN RESPONSE FACTOR 4	X	X	X	X	1.56	1	43.76	Yes	0.5	5' RACE, *ago7, rdr6, sgs3, tas3* mutant, degradome	Peragine et al., 2004; Allen et al., 2005; Hunter et al., 2006; German et al., 2008; Marin et al., 2010	UUCUUGACCUUGUAAGACCUU
TasiARFs	AT5G62000	ARF2, AUXIN RESPONSE FACTOR 2	X	X	X	X	1.852	1	629.07	Yes	0.5	*ago7*, *tas3* mutant, degradome	German et al., 2008; Marin et al., 2010	UUCUUGACCUUGUAAGACCUU

Table S2. All *A. thaliana* miRNAs retrieved from miRBase v22 and their conservation group

miRNA	miRNA sequence	miRNA family	Conservation
ath-miR156a-3p	GCUCACUGCUCUUUCUGUCAGA	miR156/miR157	Conserved
ath-miR156a-5p	UGACAGAAGAGAGUGAGCAC	miR156/miR157	Conserved
ath-miR156b-3p	UGCUCACCUCUCUUUCUGUCAGU	miR156/miR157	Conserved
ath-miR156b-5p	UGACAGAAGAGAGUGAGCAC	miR156/miR157	Conserved
ath-miR156c-3p	GCUCACUGCUCUAUCUGUCAGA	miR156/miR157	Conserved
ath-miR156c-5p	UGACAGAAGAGAGUGAGCAC	miR156/miR157	Conserved
ath-miR156d-3p	GCUCACUCUCUUUUUGUCAUAAC	miR156/miR157	Conserved
ath-miR156d-5p	UGACAGAAGAGAGUGAGCAC	miR156/miR157	Conserved
ath-miR156e	UGACAGAAGAGAGUGAGCAC	miR156/miR157	Conserved
ath-miR156f-3p	GCUCACUCUCUAUCCGUCACC	miR156/miR157	Conserved
ath-miR156f-5p	UGACAGAAGAGAGUGAGCAC	miR156/miR157	Conserved
ath-miR156g	CGACAGAAGAGAGUGAGCAC	miR156/miR157	Conserved
ath-miR156h	UGACAGAAGAAAGAGAGCAC	miR156/miR157	Conserved
ath-miR156i	UGACAGAAGAGAGAGAGCAG	miR156/miR157	Conserved
ath-miR156j	UGACAGAAGAGAGAGAGCAC	miR156/miR157	Conserved
ath-miR157a-3p	GCUCUCUAGCCUUCUGUCAUC	miR156/miR157	Conserved
ath-miR157a-5p	UUGACAGAAGAUAGAGAGCAC	miR156/miR157	Conserved
ath-miR157b-3p	GCUCUCUAGCCUUCUGUCAUC	miR156/miR157	Conserved
ath-miR157b-5p	UUGACAGAAGAUAGAGAGCAC	miR156/miR157	Conserved
ath-miR157c-3p	GCUCUCUAUACUUCUGUCACC	miR156/miR157	Conserved
ath-miR157c-5p	UUGACAGAAGAUAGAGAGCAC	miR156/miR157	Conserved
ath-miR157d	UGACAGAAGAUAGAGAGCAC	miR156/miR157	Conserved
ath-miR159a	UUUGGAUUGAAGGGAGCUCUA	miR159	Conserved
ath-miR159b-3p	UUUGGAUUGAAGGGAGCUCUU	miR159	Conserved
ath-miR159b-5p	GAGCUCCUUGAAGUUCAAUGG	miR159	Conserved
ath-miR159c	UUUGGAUUGAAGGGAGCUCCU	miR159	Conserved
ath-miR160a-3p	GCGUAUGAGGAGCCAUGCAUA	miR160	Conserved
ath-miR160a-5p	UGCCUGGCUCCCUGUAUGCCA	miR160	Conserved
ath-miR160b	UGCCUGGCUCCCUGUAUGCCA	miR160	Conserved
ath-miR160c-3p	CGUACAAGGAGUCAAGCAUGA	miR160	Conserved
ath-miR160c-5p	UGCCUGGCUCCCUGUAUGCCA	miR160	Conserved
ath-miR162a-3p	UCGAUAAACCUCUGCAUCCAG	miR162	Conserved
ath-miR162a-5p	UGGAGGCAGCGGUUCAUCGAUC	miR162	Conserved
ath-miR162b-3p	UCGAUAAACCUCUGCAUCCAG	miR162	Conserved
ath-miR162b-5p	UGGAGGCAGCGGUUCAUCGAUC	miR162	Conserved
ath-miR164a	UGGAGAAGCAGGGCACGUGCA	miR164	Conserved
ath-miR164b-3p	CAUGUGCCCAUCUUCACCAUC	miR164	Conserved

ath-miR164b-5p	UGGAGAAGCAGGGCACGUGCA	miR164	Conserved
ath-miR164c-3p	CACGUGUUCUACUACUCCAAC	miR164	Conserved
ath-miR164c-5p	UGGAGAAGCAGGGCACGUGCG	miR164	Conserved
ath-miR165a-3p	UCGGACCAGGCUUCAUCCCCC	miR165/miR166	Conserved
ath-miR165a-5p	GGAAUGUUGUCUGGAUCGAGG	miR165/miR166	Conserved
ath-miR165b	UCGGACCAGGCUUCAUCCCCC	miR165/miR166	Conserved
ath-miR166a-3p	UCGGACCAGGCUUCAUUCCCC	miR165/miR166	Conserved
ath-miR166a-5p	GGACUGUUGUCUGGCUCGAGG	miR165/miR166	Conserved
ath-miR166b-3p	UCGGACCAGGCUUCAUUCCCC	miR165/miR166	Conserved
ath-miR166b-5p	GGACUGUUGUCUGGCUCGAGG	miR165/miR166	Conserved
ath-miR166c	UCGGACCAGGCUUCAUUCCCC	miR165/miR166	Conserved
ath-miR166d	UCGGACCAGGCUUCAUUCCCC	miR165/miR166	Conserved
ath-miR166e-3p	UCGGACCAGGCUUCAUUCCCC	miR165/miR166	Conserved
ath-miR166e-5p	GGAAUGUUGUCUGGCACGAGG	miR165/miR166	Conserved
ath-miR166f	UCGGACCAGGCUUCAUUCCCC	miR165/miR166	Conserved
ath-miR166g	UCGGACCAGGCUUCAUUCCCC	miR165/miR166	Conserved
ath-miR167a-3p	GAUCAUGUUCGCAGUUUCACC	miR167	Conserved
ath-miR167a-5p	UGAAGCUGCCAGCAUGAUCUA	miR167	Conserved
ath-miR167b	UGAAGCUGCCAGCAUGAUCUA	miR167	Conserved
ath-miR167c-3p	UAGGUCAUGCUGGUAGUUUCACC	miR167	Conserved
ath-miR167c-5p	UAAGCUGCCAGCAUGAUCUUG	miR167	Conserved
ath-miR167d	UGAAGCUGCCAGCAUGAUCUGG	miR167	Conserved
ath-miR168a-3p	CCCGCCUUGCAUCAACUGAAU	miR168	Conserved
ath-miR168a-5p	UCGCUUGGUGCAGGUCGGGAA	miR168	Conserved
ath-miR168b-3p	CCCGUCUUGUAUCAACUGAAU	miR168	Conserved
ath-miR168b-5p	UCGCUUGGUGCAGGUCGGGAA	miR168	Conserved
ath-miR169a-3p	GGCAAGUUGUCCUUGGCUAC	miR169	Conserved
ath-miR169a-5p	CAGCCAAGGAUGACUUGCCGA	miR169	Conserved
ath-miR169b-3p	GGCAAGUUGUCCUUCGGCUACA	miR169	Conserved
ath-miR169b-5p	CAGCCAAGGAUGACUUGCCGG	miR169	Conserved
ath-miR169c	CAGCCAAGGAUGACUUGCCGG	miR169	Conserved
ath-miR169d	UGAGCCAAGGAUGACUUGCCG	miR169	Conserved
ath-miR169e	UGAGCCAAGGAUGACUUGCCG	miR169	Conserved
ath-miR169f-3p	GCAAGUUGACCUUGGCUCUGC	miR169	Conserved
ath-miR169f-5p	UGAGCCAAGGAUGACUUGCCG	miR169	Conserved
ath-miR169g-3p	UCCGGCAAGUUGACCUUGGCU	miR169	Conserved
ath-miR169g-5p	UGAGCCAAGGAUGACUUGCCG	miR169	Conserved
ath-miR169h	UAGCCAAGGAUGACUUGCCUG	miR169	Conserved
ath-miR169i	UAGCCAAGGAUGACUUGCCUG	miR169	Conserved
ath-miR169j	UAGCCAAGGAUGACUUGCCUG	miR169	Conserved

ath-miR169k	UAGCCAAGGAUGACUUGCCUG	miR169	Conserved
ath-miR169l	UAGCCAAGGAUGACUUGCCUG	miR169	Conserved
ath-miR169m	UAGCCAAGGAUGACUUGCCUG	miR169	Conserved
ath-miR169n	UAGCCAAGGAUGACUUGCCUG	miR169	Conserved
ath-miR170-3p	UGAUUGAGCCGUGUCAAUAUC	miR170/miR171	Conserved
ath-miR170-5p	UAUUGGCCUGGUUCACUCAGA	miR170/miR171	Conserved
ath-miR171a-3p	UGAUUGAGCCGCGCCAAUAUC	miR170/miR171	Conserved
ath-miR171a-5p	UAUUGGCCUGGUUCACUCAGA	miR170/miR171	Conserved
ath-miR171b-3p	UUGAGCCGUGCCAAUAUCACG	miR170/miR171	Conserved
ath-miR171b-5p	AGAUAUUAGUGCGGUUCAAUC	miR170/miR171	Conserved
ath-miR171c-3p	UUGAGCCGUGCCAAUAUCACG	miR170/miR171	Conserved
ath-miR171c-5p	AGAUAUUGGUGCGGUUCAAUC	miR170/miR171	Conserved
ath-miR172a	AGAAUCUUGAUGAUGCUGCAU	miR172	Conserved
ath-miR172b-3p	AGAAUCUUGAUGAUGCUGCAU	miR172	Conserved
ath-miR172b-5p	GCAGCACCAUUAAGAUUCAC	miR172	Conserved
ath-miR172c	AGAAUCUUGAUGAUGCUGCAG	miR172	Conserved
ath-miR172d-3p	AGAAUCUUGAUGAUGCUGCAG	miR172	Conserved
ath-miR172d-5p	GCAACAUCUUCAAGAUUCAGA	miR172	Conserved
ath-miR172e-3p	GGAAUCUUGAUGAUGCUGCAU	miR172	Conserved
ath-miR172e-5p	GCAGCACCAUUAAGAUUCAC	miR172	Conserved
ath-miR2111a-3p	GUCCUCGGGAUGCGGAUUACC	miR2111	Conserved
ath-miR2111a-5p	UAAUCUGCAUCCUGAGGUUUA	miR2111	Conserved
ath-miR2111b-3p	AUCCUCGGGAUACAGUUUACC	miR2111	Conserved
ath-miR2111b-5p	UAAUCUGCAUCCUGAGGUUUA	miR2111	Conserved
ath-miR319a	UUGGACUGAAGGGAGCUCCCU	miR319	Conserved
ath-miR319b	UUGGACUGAAGGGAGCUCCCU	miR319	Conserved
ath-miR319c	UUGGACUGAAGGGAGCUCCUU	miR319	Conserved
ath-miR390a-3p	CGCUAUCCAUCCUGAGUUUCA	miR390	Conserved
ath-miR390a-5p	AAGCUCAGGAGGGAUAGCGCC	miR390	Conserved
ath-miR390b-3p	CGCUAUCCAUCCUGAGUUCC	miR390	Conserved
ath-miR390b-5p	AAGCUCAGGAGGGAUAGCGCC	miR390	Conserved
ath-miR391-3p	ACGGUAUCUCUCCUACGUAGC	miR391	Conserved
ath-miR391-5p	UUCGCAGGAGAGAUAGCGCCA	miR391	Conserved
ath-miR393a-3p	AUCAUGCUAUCUCUUUGGAUU	miR393	Conserved
ath-miR393a-5p	UCCAAAGGGAUCGCAUUGAUCC	miR393	Conserved
ath-miR393b-3p	AUCAUGCGAUCUCUUUGGAUU	miR393	Conserved
ath-miR393b-5p	UCCAAAGGGAUCGCAUUGAUCC	miR393	Conserved
ath-miR394a	UUGGCAUUCUGUCCACCUCC	miR394	Conserved
ath-miR394b-3p	AGGUGGGCAUACUGCCAAUAG	miR394	Conserved
ath-miR394b-5p	UUGGCAUUCUGUCCACCUCC	miR394	Conserved

ath-miR395a	CUGAAGUGUUUGGGGGAACUC	miR395	Conserved
ath-miR395b	CUGAAGUGUUUGGGGGGACUC	miR395	Conserved
ath-miR395c	CUGAAGUGUUUGGGGGGACUC	miR395	Conserved
ath-miR395d	CUGAAGUGUUUGGGGGAACUC	miR395	Conserved
ath-miR395e	CUGAAGUGUUUGGGGGAACUC	miR395	Conserved
ath-miR395f	CUGAAGUGUUUGGGGGGACUC	miR395	Conserved
ath-miR396a-3p	GUUCAAUAAAGCUGUGGGAAG	miR396	Conserved
ath-miR396a-5p	UUCCACAGCUUUCUUGAACUG	miR396	Conserved
ath-miR396b-3p	GCUCAAGAAAGCUGUGGGAAA	miR396	Conserved
ath-miR396b-5p	UUCCACAGCUUUCUUGAACUU	miR396	Conserved
ath-miR397a	UCAUUGAGUGCAGCGUUGAUG	miR397	Conserved
ath-miR397b	UCAUUGAGUGCAUCGUUGAUG	miR397	Conserved
ath-miR398a-3p	UGUGUUCUCAGGUCACCCCUU	miR398	Conserved
ath-miR398a-5p	AAGGAGUGGCAUGUGAACACA	miR398	Conserved
ath-miR398b-3p	UGUGUUCUCAGGUCACCCCUG	miR398	Conserved
ath-miR398b-5p	AGGGUUGAUAUGAGAACACAC	miR398	Conserved
ath-miR398c-3p	UGUGUUCUCAGGUCACCCCUG	miR398	Conserved
ath-miR398c-5p	AGGGUUGAUAUGAGAACACAC	miR398	Conserved
ath-miR399a	UGCCAAAGGAGAUUUGCCCUG	miR399	Conserved
ath-miR399b	UGCCAAAGGAGAGUUGCCCUG	miR399	Conserved
ath-miR399c-3p	UGCCAAAGGAGAGUUGCCCUG	miR399	Conserved
ath-miR399c-5p	GGGCAUCUUUCUAUUGGCAGG	miR399	Conserved
ath-miR399d	UGCCAAAGGAGAUUUGCCCCG	miR399	Conserved
ath-miR399e	UGCCAAAGGAGAUUUGCCUCG	miR399	Conserved
ath-miR399f	UGCCAAAGGAGAUUUGCCCGG	miR399	Conserved
ath-miR403-3p	UUAGAUUCACGCACAAACUCG	miR403	Conserved
ath-miR403-5p	UGUUUUGUGCUUGAAUCUAAUU	miR403	Conserved
ath-miR408-3p	AUGCACUGCCUCUUCCCUGGC	miR408	Conserved
ath-miR408-5p	ACAGGGAACAAGCAGAGCAUG	miR408	Conserved
ath-miR827	UUAGAUGACCAUCAACAAACU	miR827	Conserved
ath-miR828	UCUUGCUUAAAUGAGUAUUCCA	miR828	Conserved
ath-miR845a	CGGCUCUGAUACCAAUUGAUG	miR845	Conserved
ath-miR845b	UCGCUCUGAUACCAAAUUGAUG	miR845	Conserved
ath-miR158a-3p	UCCCAAAUGUAGACAAAGCA	miR158	Brassicaceae
ath-miR158a-5p	CUUUGUCUACAAUUUUGGAAA	miR158	Brassicaceae
ath-miR158b	CCCCAAAUGUAGACAAAGCA	miR158	Brassicaceae
ath-miR161.1	UGAAAGUGACUACAUCGGGGU	miR161	Brassicaceae
ath-miR161.2	UCAAUGCAUUGAAAGUGACUA	miR161	Brassicaceae
ath-miR163	UUGAAGAGGACUUGGAACUUCGAU	miR163	Brassicaceae
ath-miR173-3p	UGAUUCUCUGUGUAAGCGAAA	miR173	Brassicaceae

ath-miR173-5p	UUCGCUUGCAGAGAGAAAUCAC	miR173	Brassicaceae
ath-miR1887	UACUAAGUAGAGUCUAAGAGA	miR1887	Brassicaceae
ath-miR2112-3p	CUUUAUAUCCGCAUUUGCGCA	miR2112	Brassicaceae
ath-miR2112-5p	CGCAAAUGCGGAUAUCAAUGU	miR2112	Brassicaceae
ath-miR3434-3p	UCAGAGUAUCAGCCAUGUGA	miR3434	Brassicaceae
ath-miR3434-5p	ACUUGGCUGAUUCUAUUAUU	miR3434	Brassicaceae
ath-miR3440b-3p	UGGAUUGGUCAAGGGAAGCGU	miR3440	Brassicaceae
ath-miR3440b-5p	UUUUCUUGGCCCAUCCACUUC	miR3440	Brassicaceae
ath-miR400	UAUGAGAGUAUUAUAAGUCAC	miR400	Brassicaceae
ath-miR402	UUCGAGGCCUAUUAAACCUCUG	miR402	Brassicaceae
ath-miR4221	UUUUCCUCUGUUGAAUUCUUGC	miR4221	Brassicaceae
ath-miR4227	UCACUGGUACCAAUCAUUCCA	miR4227	Brassicaceae
ath-miR4228-3p	UCGGAUGCGAAACGGUGGUGU	miR4228	Brassicaceae
ath-miR4228-5p	AUAGCCUUGAACGCCGUCGUU	miR4228	Brassicaceae
ath-miR4239	UUUGUUAUUUUCGCAUGCUCC	miR4239	Brassicaceae
ath-miR4240	UGACUAGACCCGUAACAUUAC	miR4240	Brassicaceae
ath-miR4243	UUGAAAUUGUAGAUUUCGUAC	miR4243	Brassicaceae
ath-miR4245	ACAAAGUUUUAUACUGACAAU	miR4245	Brassicaceae
ath-miR472-3p	UUUUUCCUACUCCGCCCAUACC	miR472	Brassicaceae
ath-miR472-5p	AUGGUCGAAGUAGGCAAAAUC	miR472	Brassicaceae
ath-miR5654-3p	UGGAAGAUGCUUUGGGAUUUAUU	miR5654	Brassicaceae
ath-miR5654-5p	AUAAAUCCCAACAUCUUCCA	miR5654	Brassicaceae
ath-miR774a	UUGGUUACCCAUAUGGCCAUC	miR774	Brassicaceae
ath-miR774b-3p	CAUCCAUAUUUUCAUCUCGAA	miR774	Brassicaceae
ath-miR774b-5p	UGAGAUGAAGAUAUGGGUGAU	miR774	Brassicaceae
ath-miR781a	UUAGAGUUUUCUGGAUACUUA	miR781	Brassicaceae
ath-miR781b	UUAGAGUUUUCUGGAUACUUA	miR781	Brassicaceae
ath-miR8171	AUAGGUGGGCCAGUGGUAGGA	miR8171	Brassicaceae
ath-miR822-3p	UGUGCAAAUGCUUUCUACAGG	miR822	Brassicaceae
ath-miR822-5p	UGCGGGAAGCAUUUGCACAUG	miR822	Brassicaceae
ath-miR823	UGGGUGGUGAUCAUAUAAGAU	miR823	Brassicaceae
ath-miR824-3p	CCUUCUCAUCGAUGGUCUAGA	miR824	Brassicaceae
ath-miR824-5p	UAGACCAUUUGUGAGAAGGGA	miR824	Brassicaceae
ath-miR825	UUCUCAAGAAGGUGCAUGAAC	miR825	Brassicaceae
ath-miR829-3p.1	AGCUCUGAUACCAAAUGAUGGAAU	miR829	Brassicaceae
ath-miR829-3p.2	CAAAUUAAAGCUUCAAGGUAG	miR829	Brassicaceae
ath-miR829-5p	ACUUUGAAGCUUUGAUUUGAA	miR829	Brassicaceae
ath-miR831-3p	UGAUCUCUUCGUACUCUUCUUG	miR831	Brassicaceae
ath-miR831-5p	AGAAGCGUACAAGGAGAUGAGG	miR831	Brassicaceae
ath-miR833a-3p	UAGACCGAUGUCAACAAACAAG	miR833	Brassicaceae

ath-miR833a-5p	UGUUUGUUGUACUCGGUCUAGU	miR833	Brassicaceae
ath-miR833b	UGUUUGUUGACAUCGGUCUAG	miR833	Brassicaceae
ath-miR834	UGGUAGCAGUAGCGGUGGUAA	miR834	Brassicaceae
ath-miR835-3p	UGGAGAAGAUACGCAAGAAAG	miR835	Brassicaceae
ath-miR835-5p	UUCUUGCAUAUGUUCUUUAUC	miR835	Brassicaceae
ath-miR837-3p	AAACGAACAAAAAACUGAUGG	miR837	Brassicaceae
ath-miR837-5p	AUCAGUUUCUUGUUCGUUUCA	miR837	Brassicaceae
ath-miR838	UUUUCUUCUACUUCUUGCACA	miR838	Brassicaceae
ath-miR839-5p	UACCAACCUUUCAUCGUUCCC	miR839	Brassicaceae
ath-miR840-3p	UUGUUUAGGUCCCUUAGUUUC	miR840	Brassicaceae
ath-miR840-5p	ACACUGAAGGACCUAAACUAAC	miR840	Brassicaceae
ath-miR841a-3p	AUUUCUAGUGGGUCGUAUUCA	miR841	Brassicaceae
ath-miR841a-5p	UACGAGCCACUUGAAACUGAA	miR841	Brassicaceae
ath-miR841b-3p	CAAUUUCUAGUGGGUCGUAUU	miR841	Brassicaceae
ath-miR841b-5p	UACGAGCCACUGGAAACUGAA	miR841	Brassicaceae
ath-miR842	UCAUGGUCAGAUCCGUCAUCC	miR842	Brassicaceae
ath-miR844-3p	UUAUAAGCCAUCUUACUAGUU	miR844	Brassicaceae
ath-miR844-5p	UGGUAAGAUUGCUUAUAAGCU	miR844	Brassicaceae
ath-miR846-3p	UUGAAUUGAAGUGCUUGAAUU	miR846	Brassicaceae
ath-miR846-5p	CAUUCAAGGACUUCUAUUCAG	miR846	Brassicaceae
ath-miR847	UCACUCCUCUUCUUCUUGAUG	miR847	Brassicaceae
ath-miR848	UGACAUGGGACUGCCUAAGCUA	miR848	Brassicaceae
ath-miR851-3p	UGGGUGGCAAACAAAGACGAC	miR851	Brassicaceae
ath-miR851-5p	UCUCGGUUCGCGAUCCACAAG	miR851	Brassicaceae
ath-miR852	AAGAUAAGCGCCUUAGUUCUG	miR852	Brassicaceae
ath-miR853	UCCCCUCUUUAGCUUGGAGAAG	miR853	Brassicaceae
ath-miR856	UAAUCCUACCAAUAACUUCAGC	miR856	Brassicaceae
ath-miR857	UUUUGUAUGUUGAAGGUGUAU	miR857	Brassicaceae
ath-miR858a	UUUCGUUGUCUGUUCGACCUU	miR858	Brassicaceae
ath-miR858b	UUCGUUGUCUGUUCGACCUUG	miR858	Brassicaceae
ath-miR859	UCUCUCUGUUGUGAAGUCAAA	miR859	Brassicaceae
ath-miR860	UCAAUAGAUUGGACUAUGUAU	miR860	Brassicaceae
ath-miR861-3p	GAUGGAUAUGUCUUCAAGGAC	miR861	Brassicaceae
ath-miR861-5p	CCUUGGAGAAAUAUGCGUCAA	miR861	Brassicaceae
ath-miR862-3p	AUAUGCUGGAUCUACUUGAAG	miR862	Brassicaceae
ath-miR862-5p	UCCAAUAGGUCGAGCAUGUGC	miR862	Brassicaceae
ath-miR868-3p	CUUCUUAAGUGCUGAUAAUGC	miR868	Brassicaceae
ath-miR868-5p	UCAUGUCGUAAUAGUAGUCAC	miR868	Brassicaceae
ath-miR869.1	AUUGGUUCAAUUCUGGUGUUG	miR869	Brassicaceae
ath-miR869.2	UCUGGUGUUGAGAUAGUUGAC	miR869	Brassicaceae

ath-miR10515	ACCCCGAUGGUUAUCCUCACC	miR10515	A_thaliana_only
ath-miR1886.1	UGAGAGAAGUGAGAUGAAAUC	miR1886	A_thaliana_only
ath-miR1886.2	UGAGAUGAAAUCUUUGAUUGG	miR1886	A_thaliana_only
ath-miR1886.3	AAUUAAAGAUUUCAUCUUACU	miR1886	A_thaliana_only
ath-miR1888a	UAAGUUAAGAUUUGUGAAGAA	miR1888	A_thaliana_only
ath-miR1888b	UUAGGCUAAGAUUUGUGAAGA	miR1888	A_thaliana_only
ath-miR2933a	GAAAUCGGAGAGGAAAUUCGCC	miR2933	A_thaliana_only
ath-miR2933b	GAAAUCGGAGAGGAAAUUCGCC	miR2933	A_thaliana_only
ath-miR2934-3p	CAUCCAAGGUGUUUGUAGAAA	miR2934	A_thaliana_only
ath-miR2934-5p	UCUUUCUGCAAACGCCUUGGA	miR2934	A_thaliana_only
ath-miR2936	CUUGAGAGAGAGAACACAGACG	miR2936	A_thaliana_only
ath-miR2937	AUAAGAGCUGUUGAAGGAGUC	miR2937	A_thaliana_only
ath-miR2938	GAUCUUUUGAGAGGGUUCCAG	miR2938	A_thaliana_only
ath-miR2939	UAACGCACAACACUAAGCCAU	miR2939	A_thaliana_only
ath-miR3932a	AACUUUGUGAUGACAACGAAG	miR3932	A_thaliana_only
ath-miR3932b-3p	AACUUUGUGAUGACAACGAAG	miR3932	A_thaliana_only
ath-miR3932b-5p	UUUGACGUGCUCGAUCUGCUC	miR3932	A_thaliana_only
ath-miR3933	AGAAGCAAAAUGACGACUCGG	miR3933	A_thaliana_only
ath-miR401	CGAAACUGGUGUCGACCGACA	miR401	A_thaliana_only
ath-miR404	AUUAACGCUGGCGGUUGCGGCAGC	miR404	A_thaliana_only
ath-miR405a	AUGAGUUGGGUCUAACCCAUAACU	miR405	A_thaliana_only
ath-miR405b	AUGAGUUGGGUCUAACCCAUAACU	miR405	A_thaliana_only
ath-miR405d	AUGAGUUGGGUCUAACCCAUAACU	miR405	A_thaliana_only
ath-miR406	UAGAAUGCUAUUGUAAUCCAG	miR406	A_thaliana_only
ath-miR407	UUUAAAUCAUAUACUUUUGGU	miR407	A_thaliana_only
ath-miR413	AUAGUUUCUCUUGUUCUGCAC	miR413	A_thaliana_only
ath-miR414	UCAUCUUCAUCAUCAUCGUCA	miR414	A_thaliana_only
ath-miR415	AACAGAGCAGAAACAGAACAU	miR415	A_thaliana_only
ath-miR416	GGUUCGUACGUACACUGUUCA	miR416	A_thaliana_only
ath-miR417	GAAGGUAGUGAAUUUGUUCGA	miR417	A_thaliana_only
ath-miR418	UAAUGUGAUGAUGAACUGACC	miR418	A_thaliana_only
ath-miR419	UUAUGAAUGCUGAGGAUGUUG	miR419	A_thaliana_only
ath-miR420	UAAACUAAUCACGGAAAUGCA	miR420	A_thaliana_only
ath-miR426	UUUUGGAAAUUUGUCCUUACG	miR426	A_thaliana_only
ath-miR447a.2-3p	UAUGGAAGAAAUUGUAGUAUU	miR447	A_thaliana_only
ath-miR447a-3p	UUGGGGACGAGAUGUUUUGUUG	miR447	A_thaliana_only
ath-miR447b	UUGGGGACGAGAUGUUUUGUUG	miR447	A_thaliana_only
ath-miR447c-3p	UUGGGGACGACAUCUUUUGUUG	miR447	A_thaliana_only
ath-miR447c-5p	CCCCUUACAAUGUCGAGUAAA	miR447	A_thaliana_only
ath-miR5012	UUUUACUGCUACUUGUGUUCC	miR5012	A_thaliana_only

ath-miR5013	UUUGUGACAUCUAGGUGCUUU	miR5013	A_thaliana_only
ath-miR5014a-3p	UUGUACAAAUUUAAGUGUACG	miR5014	A_thaliana_only
ath-miR5014a-5p	ACACUUAGUUUUGUACAACAU	miR5014	A_thaliana_only
ath-miR5014b	AUUUGUACACCUAGAUCUGUA	miR5014	A_thaliana_only
ath-miR5015	UUGGUGUUAUGUGUAGUCUUC	miR5015	A_thaliana_only
ath-miR5016	UUCUUGUGGAUUCCUUGGAAA	miR5016	A_thaliana_only
ath-miR5017-3p	UUAUACCAAAUUAAUAGCAAA	miR5017	A_thaliana_only
ath-miR5017-5p	AUUUGUUACUAAUUUGGAAUG	miR5017	A_thaliana_only
ath-miR5018	UUAAAGCUCCACCAUGAGUCCAAU	miR5018	A_thaliana_only
ath-miR5019	UGUUGGGAAAGAAAAACUCUU	miR5019	A_thaliana_only
ath-miR5020a	UGGAAGAAGGUGAGACUUGCA	miR5020	A_thaliana_only
ath-miR5020b	AUGGCAUGAAAGAAGGUGAGA	miR5020	A_thaliana_only
ath-miR5020c	UGGCAUGGAAGAAGGUGAGAC	miR5020	A_thaliana_only
ath-miR5021	UGAGAAGAAGAAGAAGAAAA	miR5021	A_thaliana_only
ath-miR5022	GUCAUGGGGUAUGAUCGAAUG	miR5022	A_thaliana_only
ath-miR5023	AUUGGUAGUGGAUAAGGGGGC	miR5023	A_thaliana_only
ath-miR5024-3p	CCGUAUCUUGGCCUUGUCAUU	miR5024	A_thaliana_only
ath-miR5024-5p	AUGACAAGGCCAAGAUAUAACA	miR5024	A_thaliana_only
ath-miR5025	ACUGUAUAUAUGUAAGUGACA	miR5025	A_thaliana_only
ath-miR5026	ACUCAUAAGAUCGUGACACGU	miR5026	A_thaliana_only
ath-miR5027	ACCGGUUGGAACUUGCCUUAA	miR5027	A_thaliana_only
ath-miR5028	AAUUGGGUUUAUGCUAGAGUU	miR5028	A_thaliana_only
ath-miR5029	AAUGAGAGAGAACACUGCAAA	miR5029	A_thaliana_only
ath-miR5595a	ACAUAUGAUCUGCAUCUUUGC	miR5595	A_thaliana_only
ath-miR5628	GAAAUAGCGAAGAUAUGAUUA	miR5628	A_thaliana_only
ath-miR5629	UUAGGGUAGUUAACGGAAGUUA	miR5629	A_thaliana_only
ath-miR5630a	GCUAAGAGCGGUUCUGAUGGA	miR5630	A_thaliana_only
ath-miR5630b	GCUAAGAGCGGUUCUGAUGGA	miR5630	A_thaliana_only
ath-miR5631	UGGCAGGAAAGACAUAAUUUU	miR5631	A_thaliana_only
ath-miR5632-3p	UUGGAUUUAUAGUUGGAUAAG	miR5632	A_thaliana_only
ath-miR5632-5p	UUGAUUCUCUUAUCCAACUGU	miR5632	A_thaliana_only
ath-miR5633	UAUGAUCAUCAGAAAACAGUG	miR5633	A_thaliana_only
ath-miR5634	AGGGACUUUGUGAAUUUAGGG	miR5634	A_thaliana_only
ath-miR5635a	UGUUAAGGAGUGUUAACGGUG	miR5635	A_thaliana_only
ath-miR5635b	UGUUAAGGAGUGUUAACGGUG	miR5635	A_thaliana_only
ath-miR5635c	UGUUAAGGAGUGUUAACGGUG	miR5635	A_thaliana_only
ath-miR5635d	UGUUAAGGAGUGUUAACGGUG	miR5635	A_thaliana_only
ath-miR5636	CGUAGUUGCAGAGCUUGACGG	miR5636	A_thaliana_only
ath-miR5637	AAUGCGCAACUCUAUAUUUCC	miR5637	A_thaliana_only
ath-miR5638a	AUACCAAAACUCUCUCACUUU	miR5638	A_thaliana_only

ath-miR5638b	ACAGUGGUCAUCUGGUGGGCU	miR5638	A_thaliana_only
ath-miR5639-3p	UUUAGCCUCAGACCACGGUGGACU	miR5639	A_thaliana_only
ath-miR5639-5p	UAGUCCACUGUGGUCUAAGGC	miR5639	A_thaliana_only
ath-miR5640	UGAGAGAAGGAAUUAGAUUCA	miR5640	A_thaliana_only
ath-miR5641	UGGAAGAAGAUGAUAGAAUUA	miR5641	A_thaliana_only
ath-miR5642a	UCUCGCGCUUGUACGGCUUU	miR5642	A_thaliana_only
ath-miR5642b	UCUCGCGCUUGUACGGCUUU	miR5642	A_thaliana_only
ath-miR5643a	AGGCUUUUAAGAUCUGGUUGC	miR5643	A_thaliana_only
ath-miR5643b	AGGCUUUUAAGAUCUGGUUGC	miR5643	A_thaliana_only
ath-miR5644	GUGGGUUGCGGAUAACGGUA	miR5644	A_thaliana_only
ath-miR5645a	AUUUGAGUCAUGUCGUUAAG	miR5645	A_thaliana_only
ath-miR5645b	AUUUGAGUCAUGUCGUUAAG	miR5645	A_thaliana_only
ath-miR5645c	AACCUAUUUAACGACAUGACU	miR5645	A_thaliana_only
ath-miR5645d	AUUUGAGUCAUGUCGUUAAG	miR5645	A_thaliana_only
ath-miR5645e	AUUUGAGUCAUGUCGUUAAG	miR5645	A_thaliana_only
ath-miR5645f	AUUUGAGUCAUGUCGUUAAG	miR5645	A_thaliana_only
ath-miR5646	GUUCGAGGCACGUUGGGAGG	miR5646	A_thaliana_only
ath-miR5647	UCAAGUUUGAUGACGAUUCCA	miR5647	A_thaliana_only
ath-miR5648-3p	AUCUGAAGAAAAUAGCGGCAU	miR5648	A_thaliana_only
ath-miR5648-5p	UUUGGAAAUAUUUGGCUUGACU	miR5648	A_thaliana_only
ath-miR5649a	AUUGAAUAUGUUGGUUACUAU	miR5649	A_thaliana_only
ath-miR5649b	AUUGAAUAUGUUGGUUACUAU	miR5649	A_thaliana_only
ath-miR5650	UUGUUUUGGAUCUUAGAUACA	miR5650	A_thaliana_only
ath-miR5651	UUGUGCGGUUCAAAUAGUAAC	miR5651	A_thaliana_only
ath-miR5652	UUGAAUGUGAAUGAAUCGGGC	miR5652	A_thaliana_only
ath-miR5653	UGGGUUGAGUUGAGUUGAGUUGGC	miR5653	A_thaliana_only
ath-miR5655	AAGUAGACACAUAAGAAGGAG	miR5655	A_thaliana_only
ath-miR5656	ACUGAAGUAGAGAUUGGGUUU	miR5656	A_thaliana_only
ath-miR5657	UGGACAAGGUUAGAUUUGGUG	miR5657	A_thaliana_only
ath-miR5658	AUGAUGAUGAUGAUGAUGAAA	miR5658	A_thaliana_only
ath-miR5659	CGAUGAAGGUCUUUGGAACGGUA	miR5659	A_thaliana_only
ath-miR5660	CAGGUGGUUAGUGCAAUGGAA	miR5660	A_thaliana_only
ath-miR5661	AGAGGUACAUCAUGUAGUCUG	miR5661	A_thaliana_only
ath-miR5662	AGAGGUGACCAUUGGAGAUG	miR5662	A_thaliana_only
ath-miR5663-3p	UGAGAAUGCAAAUCCUUAGCU	miR5663	A_thaliana_only
ath-miR5663-5p	AGCUAAGGAUUUGCAUUCUCA	miR5663	A_thaliana_only
ath-miR5664	AUAGUCAAUUUUAUCGGUCUG	miR5664	A_thaliana_only
ath-miR5665	UUGGUGGACAAGAUCUGGGAU	miR5665	A_thaliana_only
ath-miR5666	AUGGGACAUCGAGCAUUUAAU	miR5666	A_thaliana_only
ath-miR5995b	ACAUAUGAUCUGCAUCUUUGC	miR5995	A_thaliana_only

ath-miR5996	UGACAUCCAGAUAGAAGCUUUG	miR5996	A_thaliana_only
ath-miR5997	UGAAACCAAGUAGCUAAAUAG	miR5997	A_thaliana_only
ath-miR5998a	ACAGUUUGUGUUUUGUUUUGU	miR5998	A_thaliana_only
ath-miR5998b	ACAGUUUGUGUUUUGUUUUGU	miR5998	A_thaliana_only
ath-miR5999	UCUUCACUAUUAGACGGACAA	miR5999	A_thaliana_only
ath-miR771	UGAGCCUCUGUGGUAGCCCUCA	miR771	A_thaliana_only
ath-miR773a	UUUGCUUCCAGCUUUUGUCUC	miR773	A_thaliana_only
ath-miR773b-3p	UUUGAUUCCAGCUUUUGUCUC	miR773	A_thaliana_only
ath-miR773b-5p	GGCAAUAACUUGAGCAAACA	miR773	A_thaliana_only
ath-miR775	UUCGAUGUCUAGCAGUGCCA	miR775	A_thaliana_only
ath-miR776	UCUAAGUCUUCUAUUGAUGUU	miR776	A_thaliana_only
ath-miR777	UACGCAUUGAGUUUCGUUGCUU	miR777	A_thaliana_only
ath-miR778	UGGCUUGGUUUAUGUACACCG	miR778	A_thaliana_only
ath-miR779.1	UUCUGCUAUGUUGCUGCUCAU	miR779	A_thaliana_only
ath-miR779.2	UGAUUGGAAAUUUCGUUGACU	miR779	A_thaliana_only
ath-miR780.1	UCUAGCAGCUGUUGAGCAGGU	miR780	A_thaliana_only
ath-miR780.2	UUCUUCGUGAAUAUCUGGCAU	miR780	A_thaliana_only
ath-miR782	ACAAACACCUUGGAUGUUCUU	miR782	A_thaliana_only
ath-miR8121	AAAGUAUAAUGGUUUAGUGGUUUG	miR8121	A_thaliana_only
ath-miR8165	AAUGGAGGCAAGUGUGAAGGA	miR8165	A_thaliana_only
ath-miR8166	AGAGAGUGUAGAAAGUUUCUCA	miR8166	A_thaliana_only
ath-miR8167a	AGAUGUGGAGAUCGUGGGGAUG	miR8167	A_thaliana_only
ath-miR8167b	AGAUGUGGAGAUCGUGGGGAUG	miR8167	A_thaliana_only
ath-miR8167c	AGAUGUGGAGAUCGUGGGGAUG	miR8167	A_thaliana_only
ath-miR8167d	AGAUGUGGAGAUCGUGGGGAUG	miR8167	A_thaliana_only
ath-miR8167e	AGAUGUGGAGAUCGUGGGGAUG	miR8167	A_thaliana_only
ath-miR8167f	AGAUGUGGAGAUCGUGGGGAUG	miR8167	A_thaliana_only
ath-miR8168	AGGUGCUGAGUGUGCUAGUGC	miR8168	A_thaliana_only
ath-miR8169	AUAGACAGAGUCACUCACAGA	miR8169	A_thaliana_only
ath-miR8170-3p	UUGCUUAAAGAUUUUCUAUGU	miR8170	A_thaliana_only
ath-miR8170-5p	AUAGCAAAUCGAUAAGCAAUG	miR8170	A_thaliana_only
ath-miR8172	AUGGAUCAUCUAGAUGGAGAU	miR8172	A_thaliana_only
ath-miR8173	AUGUGCUGAUUCGAGGUGGGA	miR8173	A_thaliana_only
ath-miR8174	AUGUGUAUAGGGAAGCUAAUC	miR8174	A_thaliana_only
ath-miR8175	GAUCCCCGGCAACGGCGCCA	miR8175	A_thaliana_only
ath-miR8176	GGCCGGUGGUCGCGAGAGGGA	miR8176	A_thaliana_only
ath-miR8177	GUGUGAUGAUGUGUCAUUUAUA	miR8177	A_thaliana_only
ath-miR8178	UAACAGAGUAAUUGUACAGUG	miR8178	A_thaliana_only
ath-miR8179	UGACUGCAUUAACUUGAUCGU	miR8179	A_thaliana_only
ath-miR8180	UGCGGUGCGGGAGAAGUGC	miR8180	A_thaliana_only

ath-miR8181	UGGGGGUGGGGGGGUGACAG	miR8181	A_thaliana_only
ath-miR8182	UUGUGUUGCGUUUCUGUUGAUU	miR8182	A_thaliana_only
ath-miR8183	UUUAGUUGACGGAAUUGUGGC	miR8183	A_thaliana_only
ath-miR8184	UUUGGUCUGAUUACGAAUGUA	miR8184	A_thaliana_only
ath-miR826a	UAGUCCGGUUUUGGAUACGUG	miR826	A_thaliana_only
ath-miR826b	UGGUUUUGGACACGUGAAAAU	miR826	A_thaliana_only
ath-miR830-3p	UAACUAUUUUGAGAAGAAGUG	miR830	A_thaliana_only
ath-miR830-5p	UCUUCUCCAAAUAGUUUAGGUU	miR830	A_thaliana_only
ath-miR832-3p	UUGAUUCCCAAUCCAAGCAAG	miR832	A_thaliana_only
ath-miR832-5p	UGCUGGGAUCGGGAAUCGAAA	miR832	A_thaliana_only
ath-miR836	UCCUGUGUUUCCUUUGAUGCGUGG	miR836	A_thaliana_only
ath-miR843	UUUAGGUCGAGCUUCAUUGGA	miR843	A_thaliana_only
ath-miR849	UAACUAAACAUUGGUGUAGUA	miR849	A_thaliana_only
ath-miR850	UAAGAUCCGGACUACAACAAAG	miR850	A_thaliana_only
ath-miR854a	GAUGAGGAUAGGGAGGAGGAG	miR854	A_thaliana_only
ath-miR854b	GAUGAGGAUAGGGAGGAGGAG	miR854	A_thaliana_only
ath-miR854c	GAUGAGGAUAGGGAGGAGGAG	miR854	A_thaliana_only
ath-miR854d	GAUGAGGAUAGGGAGGAGGAG	miR854	A_thaliana_only
ath-miR854e	GAUGAGGAUAGGGAGGAGGAG	miR854	A_thaliana_only
ath-miR855	AGCAAAAGCUAAGGAAAAGGAA	miR855	A_thaliana_only
ath-miR863-3p	UUGAGAGCAACAAGACAUAAU	miR863	A_thaliana_only
ath-miR863-5p	UUAUGUCUUGUUGAUCUCAAU	miR863	A_thaliana_only
ath-miR864-3p	UAAAGUCAAUAAUACCUUGAAG	miR864	A_thaliana_only
ath-miR864-5p	UCAGGUAUGAUUGACUUCAAA	miR864	A_thaliana_only
ath-miR865-3p	UUUUUCCUCAAAUUUAUCCAA	miR865	A_thaliana_only
ath-miR865-5p	AUGAAUUUGGAUCUAAUUGAG	miR865	A_thaliana_only
ath-miR866-3p	ACAAAAUCCGUCUUUGAAGA	miR866	A_thaliana_only
ath-miR866-5p	UCAAGGAACGGAUUUUGUUAA	miR866	A_thaliana_only
ath-miR867	UUGAACAUGGUUUAUUAGGAA	miR867	A_thaliana_only
ath-miR870-3p	UAAUUUGGUGUUUCUUCGAUC	miR870	A_thaliana_only
ath-miR870-5p	AAGAACAUCAAAUUAGAAUGU	miR870	A_thaliana_only

Table S3. All HE targets of conserved passenger strand miRNAs and their Category Scores

*The highest cleavage tag abundance found for this gene across all degradome libraries

miRNA	miRNA family	Gene ID	Cat score	Maximum Category	Cleavage Tag Abundance*	miRNA sequence	Strand	Gene symbol	Gene brief description
ath-miR172b-5p	miR172	AT1G23490	0.118	Cat_2	11.29	GCAGCACCAUUAAGAUUCAC	Passenger	ARF1	ADP-ribosylation factor 1
ath-miR156c-3p	miR156/ miR157	AT1G63010	0.118	Cat_2	11.01	GCUCACUGCUCUAUCUGUCAGA	Passenger		Major Facilitator Superfamily with SPX (SYG1/Pho81/XPR1) domain-containing protein
ath-miR408-5p	miR408	AT2G26250	0.118	Cat_2	65.63	ACAGGGAACAAGCAGAGCAUG	Passenger	KCS10	3-ketoacyl-CoA synthase 10
ath-miR408-5p	miR408	AT3G20920	0.118	Cat_2	15.2	ACAGGGAACAAGCAGAGCAUG	Passenger		translocation protein-related
ath-miR160c-3p	miR160	AT4G32340	0.118	Cat_2	42.49	CGUACAAGGAGUCAAGCAUGA	Passenger		Tetratricopeptide repeat (TPR)-like superfamily protein
ath-miR167c-3p	miR167	AT5G24770	0.118	Cat_2	9.4	UAGGUCAUGCUGGUAGUUUCACC	Passenger	VSP2	vegetative storage protein 2
ath-miR398b-5p	miR398	AT5G64470	0.118	Cat_2	14.57	AGGGUUGAUAUGAGAACACAC	Passenger	TBL12	Plant protein of unknown function (DUF828)
ath-miR403-5p	miR403	AT3G61050	0.147	Cat_2	13.99	UGUUUUGUGCUUGAAUCUAAUU	Passenger	NTMC2 T4	Calcium-dependent lipid-binding (CaLB domain) family protein
ath-miR398a-5p	miR398	AT4G27130	0.176	Cat_2	14.58	AAGGAGUGGCAUGUGAACACA	Passenger		Translation initiation factor SUI1 family protein
ath-miR165a-5p	miR165/ miR166	AT5G24780	0.176	Cat_2	32.88	GGAAUGUUGUCUGGAUCGAGG	Passenger	VSP1	vegetative storage protein 1
ath-miR165a-5p	miR165/ miR166	AT5G67400	0.176	Cat_2	24.75	GGAAUGUUGUCUGGAUCGAGG	Passenger	RHS19	root hair specific 19
ath-miR408-5p	miR408	AT1G32080	0.206	Cat_2	18.73	ACAGGGAACAAGCAGAGCAUG	Passenger		membrane protein, putative
ath-miR398b-5p	miR398	AT1G21460	0.235	Cat_2	23.01	AGGGUUGAUAUGAGAACACAC	Passenger		Nodulin MtN3 family protein
ath-miR398a-5p	miR398	AT4G12600	0.235	Cat_2	19.43	AAGGAGUGGCAUGUGAACACA	Passenger		Ribosomal protein L7Ae/L30e/S12e/Gadd45 family protein
ath-miR396b-3p	miR396	AT1G43170	0.265	Cat_2	28.9	GCUCAAGAAAGCUGUGGGAAA	Passenger	RP1	ribosomal protein 1
ath-miR168b-3p	miR168	AT2G43710	0.265	Cat_2	8.84	CCCGUCUUGUAUCAACUGAAU	Passenger	SSI2	Plant stearoyl-acyl-carrier-protein desaturase family protein
ath-miR408-5p	miR408	AT2G47400	0.265	Cat_2	51.04	ACAGGGAACAAGCAGAGCAUG	Passenger	CP12-1	CP12 domain-containing protein 1
ath-miR408-5p	miR408	AT4G38680	0.324	Cat_2	24.98	ACAGGGAACAAGCAGAGCAUG	Passenger	GRP2	glycine rich protein 2
ath-miR165a-5p	miR165/ miR166	AT3G51430	0.382	Cat_1	8.84	GGAAUGUUGUCUGGAUCGAGG	Passenger	YLS2	Calcium-dependent phosphotriesterase superfamily protein
ath-miR408-5p	miR408	AT1G06680	0.412	Cat_2	40.4	ACAGGGAACAAGCAGAGCAUG	Passenger	PSBP-1	photosystem II subunit P-1
ath-miR160c-3p	miR160	AT3G23810	0.471	Cat_2	55.97	CGUACAAGGAGUCAAGCAUGA	Passenger	SAHH2	S-adenosyl-l-homocysteine (SAH) hydrolase 2
ath-miR172d-5p	miR172	AT1G07320	0.5	Cat_2	97.96	GCAACAUCUUCAAGAUUCAGA	Passenger	RPL4	ribosomal protein L4
ath-miR167a-3p	miR167	AT3G03780	0.5	Cat_2	83.05	GAUCAUGUUCGCAGUUUCACC	Passenger	MS2	methionine synthase 2
ath-miR172d-5p	miR172	AT3G23810	0.5	Cat_2	164.28	GCAACAUCUUCAAGAUUCAGA	Passenger	SAHH2	S-adenosyl-l-homocysteine (SAH) hydrolase 2
ath-miR165a-5p	miR165/ miR166	AT3G51420	0.5	Cat_1	18.41	GGAAUGUUGUCUGGAUCGAGG	Passenger	SSL4	strictosidine synthase-like 4
ath-miR398b-5p	miR398	AT2G02100	0.647	Cat_2	517.73	AGGGUUGAUAUGAGAACACAC	Passenger	LCR69	low-molecular-weight cysteine-rich 69
ath-miR398a-5p	miR398	AT5G66570	0.676	Cat_2	290.82	AAGGAGUGGCAUGUGAACACA	Passenger	PSBO1	PS II oxygen-evolving complex 1

ath-miR396a-3p	miR396	No HE Targets	NA	NA	NA	GUUCAAUAAAGCUGUGGGAAG	Passenger	NA	NA
ath-miR156a-3p	miR156/ miR157	No HE Targets	NA	NA	NA	GCUCACUGCUCUUUCUGUCAGA	Passenger	NA	NA
ath-miR159b-5p	miR159	No HE Targets	NA	NA	NA	GAGCUCCUUGAAGUUCAAUGG	Passenger	NA	NA
ath-miR160a-3p	miR160	No HE Targets	NA	NA	NA	GCGUAUGAGGAGCCAUGCAUA	Passenger	NA	NA
ath-miR162a-5p	miR162	No HE Targets	NA	NA	NA	UGGAGGCAGCGGUUCAUCGAUC	Passenger	NA	NA
ath-miR164b-3p	miR164	No HE Targets	NA	NA	NA	CAUGUGCCCAUCUUCACCAUC	Passenger	NA	NA
ath-miR166e-5p	miR165/ miR166	No HE Targets	NA	NA	NA	GGAAUGUUGUCUGGCACGAGG	Passenger	NA	NA
ath-miR168a-3p	miR168	No HE Targets	NA	NA	NA	CCCGCCUUGCAUCAACUGAAU	Passenger	NA	NA
ath-miR169a-3p	miR169	No HE Targets	NA	NA	NA	GGCAAGUUGUCCUUGGCUAC	Passenger	NA	NA
ath-miR170-5p	miR170/ miR171	No HE Targets	NA	NA	NA	UAUUGGCCUGGUUCACUCAGA	Passenger	NA	NA
ath-miR390a-3p	miR390	No HE Targets	NA	NA	NA	CGCUAUCCAUCCUGAGUUUCA	Passenger	NA	NA
ath-miR393a-3p	miR393	No HE Targets	NA	NA	NA	AUCAUGCUAUCUCUUUGGAUU	Passenger	NA	NA
ath-miR394b-3p	miR394	No HE Targets	NA	NA	NA	AGGUGGGCAUACUGCCAAUAG	Passenger	NA	NA
ath-miR399c-5p	miR399	No HE Targets	NA	NA	NA	GGGCAUCUUUCUAUUGGCAGG	Passenger	NA	NA

Table S4. All HE targets of conserved guide strand miRNAs and their Category Scores

Conserved targets are highlighted in green

Non-conserved targets are highlighted in yellow

*The highest cleavage tag abundance found for this gene across all degradome libraries

miRNA	miRNA family	Gene ID	Cat score	Maximum Category	Cleavage tag abundance*	miRNA sequence	Strand	Gene symbol	Gene brief description	PANTHER ID
ath-miR2111b-3p	miR2111	AT3G57410	0.118	Cat_2	13.43	AUCCUCGGGAUACAGUUUACC	Guide	VLN3	villin 3	PTHR11977
ath-miR390a-5p	miR390	AT1G10490	0.118	Cat_2	40.66	AAGCUCAGGAGGGAUAGCGCC	Guide		Domain of unknown function (DUF1726) ;Putative ATPase (DUF699)	PTHR10925
ath-miR396a-5p	miR396	AT5G53660	0.118	Cat_2	30.06	UUCCACAGCUUUCUUGAACUG	Guide	GRF7	growth-regulating factor 7	PTHR31602
ath-miR408-3p	miR408	AT5G25630	0.118	Cat_2	13.99	AUGCACUGCCUCUUCCCUGGC	Guide		Tetratricopeptide repeat (TPR)-like superfamily protein	PTHR24015
ath-miR319a	miR319	AT2G21600	0.147	Cat_2	53.71	UUGGACUGAAGGGAGCUCCCU	Guide	RER1B	endoplasmatic reticulum retrieval protein 1B	PTHR10743
ath-miR827	miR827	AT1G33140	0.147	Cat_2	12.99	UUAGAUGACCAUCAACAAACU	Guide	PGY2	Ribosomal protein L6 family	PTHR11655
ath-miR156g	miR156/ miR157	AT2G47590	0.176	Cat_2	20.28	CGACAGAAGAGAGUGAGCAC	Guide	PHR2	photolyase/blue-light receptor 2	PTHR11455
ath-miR157d	miR156/ miR157	AT4G28660	0.176	Cat_2	18.17	UGACAGAAGAUAGAGAGCAC	Guide	PSB28	photosystem II reaction center PSB28 protein	PTHR34963
ath-miR169d	miR169	AT5G38030	0.176	Cat_2	13.21	UGAGCCAAGGAUGACUUGCCG	Guide		MATE efflux family protein	PTHR11206
ath-miR172c	miR172	AT5G35360	0.176	Cat_2	10.61	AGAAUCUUGAUGAUGCUGCAG	Guide	CAC2	acetyl Co-enzyme a carboxylase biotin carboxylase subunit	PTHR18866
ath-miR408-3p	miR408	AT2G47900	0.176	Cat_2	9.11	AUGCACUGCCUCUUCCCUGGC	Guide	TLP3	tubby like protein 3	PTHR16517
ath-miR167c-5p	miR167	AT3G16470	0.206	Cat_2	37.18	UAAGCUGCCAGCAUGAUCUUG	Guide	JR1	Mannose-binding lectin superfamily protein	PTHR23244
ath-miR170-3p	miR170/ miR171	AT1G22640	0.206	Cat_2	28.7	UGAUUGAGCCGUGUCAAUAUC	Guide	MYB3	myb domain protein 3	PTHR10641
ath-miR172a	miR172	AT3G05530	0.206	Cat_2	13.93	AGAAUCUUGAUGAUGCUGCAU	Guide	RPT5A	regulatory particle triple-A ATPase 5A	PTHR23073
ath-miR398b-3p	miR398	AT2G40400	0.206	Cat_2	20.2	UGUGUUCUCAGGUCACCCCUG	Guide		Protein of unknown function (DUF399 and DUF3411)	PTHR31620
ath-miR159a	miR159	AT2G21600	0.235	Cat_1	36.46	UUUGGAUUGAAGGGAGCUCU A	Guide	RER1B	endoplasmatic reticulum retrieval protein 1B	PTHR10743
ath-miR2111a-5p	miR2111	AT3G27150	0.235	Cat_1	6.6	UAAUCUGCAUCCUGAGGUUUA	Guide		Galactose oxidase/kelch repeat superfamily protein	PTHR24413
ath-miR390a-5p	miR390	AT5G48480	0.235	Cat_2	145.9	AAGCUCAGGAGGGAUAGCGCC	Guide		Lactoylglutathione lyase / glyoxalase I family protein	PTHR34109
ath-miR408-3p	miR408	AT2G30210	0.235	Cat_2	49.87	AUGCACUGCCUCUUCCCUGGC	Guide	LAC3	laccase 3	PTHR11709
ath-miR319a	miR319	AT2G31070	0.265	Cat_1	106.76	UUGGACUGAAGGGAGCUCCCU	Guide	TCP10	TCP domain protein 10	PTHR31072
ath-miR395a	miR395	AT5G10180	0.265	Cat_1	20.78	CUGAAGUGUUUGGGGGAACUC	Guide	SULTR2;1	slufate transporter 2;1	PTHR11814
ath-miR398a-3p	miR398	AT2G46250	0.265	Cat_1	17.08	UGUGUUCUCAGGUCACCCCUU	Guide		myosin heavy chain-related	PTHR31071
ath-miR399a	miR399	AT2G33770	0.265	Cat_2	35.35	UGCCAAAGGAGAUUUGCCCUG	Guide	PHO2	phosphate 2	PTHR24067

ath-miR408-3p	miR408	AT1G68010	0.265	Cat_2	45.37	AUGCACUGCCUCUUCCCUGGC	Guide	HPR	hydroxypyruvate reductase	PTHR10996
ath-miR398a-3p	miR398	AT1G03630	0.294	Cat_2	49.75	UGUGUUCUCAGGUCACCCCUU	Guide	POR C	protochlorophyllide oxidoreductase C	PTHR24322
ath-miR169a-5p	miR169	AT1G54160	0.324	Cat_2	50.73	CAGCCAAGGAUGACUUGCCGA	Guide	NF-YA5	nuclear factor Y, subunit A5	PTHR12632
ath-miR170-3p	miR170/miR171	AT4G18030	0.353	Cat_2	30.53	UGAUUGAGCCGUGUCAAUAUC	Guide		S-adenosyl-L-methionine-dependent methyltransferases superfamily protein	PTHR10108
ath-miR390a-5p	miR390	AT1G14510	0.353	Cat_1	12.48	AAGCUCAGGAGGGAUAGCGCC	Guide	AL7	alfin-like 7	PTHR12321
ath-miR394a	miR394	AT1G27340	0.353	Cat_1	16.25	UUGGCAUUCUGUCCACCUCC	Guide		Galactose oxidase/kelch repeat superfamily protein	PTHR32133
ath-miR395a	miR395	AT5G13630	0.353	Cat_2	72.7	CUGAAGUGUUUGGGGGAACUC	Guide	GUN5	magnesium-chelatase subunit chlH, chloroplast, putative / Mg-protoporphyrin IX chelatase, putative (CHLH)	PTHR23304
ath-miR396a-5p	miR396	AT3G19400	0.353	Cat_1	26.37	UUCCACAGCUUUCUUGAACUG	Guide		Cysteine proteinases superfamily protein	PTHR12411
ath-miR398a-3p	miR398	AT1G14700	0.353	Cat_1	9.2	UGUGUUCUCAGGUCACCCCUU	Guide	PAP3	purple acid phosphatase 3	PTHR10161
ath-miR396a-5p	miR396	AT1G60140	0.382	Cat_1	8.56	UUCCACAGCUUUCUUGAACUG	Guide	TPS10	trehalose phosphate synthase	PTHR10788
ath-miR319a	miR319	AT4G18390	0.412	Cat_1	253.67	UUGGACUGAAGGGAGCUCCCU	Guide	TCP2	TEOSINTE BRANCHED 1, cycloidea and PCF transcription factor 2	PTHR31072
ath-miR156h	miR156/miR157	AT1G15690	0.441	Cat_2	43.76	UGACAGAAGAAAGAGAGCAC	Guide	AVP1	Inorganic H pyrophosphatase family protein	PTHR31998
ath-miR159a	miR159	AT2G34010	0.471	Cat_1	51.67	UUUGGAUUGAAGGGAGCUCUA	Guide		verprolin, TCP INTERACTOR CONTAINING EAR MOTIF PROTEIN 4, TIE4	PTHR33388
ath-miR169a-5p	miR169	AT1G48500	0.471	Cat_1	9	CAGCCAAGGAUGACUUGCCGA	Guide	JAZ4	jasmonate-zim-domain protein 4	PTHR33077
ath-miR393a-5p	miR393	AT3G23690	0.471	Cat_1	11.32	UCCAAAGGGAUCGCAUUGAUCC	Guide		basic helix-loop-helix (bHLH) DNA-binding superfamily protein	PTHR12565
ath-miR172a	miR172	AT2G39250	0.5	Cat_1	18.76	AGAAUCUUGAUGAUGCUGCAU	Guide	SNZ	AP2-LIKE ETHYLENE-RESPONSIVE TRANSCRIPTION FACTOR SMZ-RELATED	PTHR32467
ath-miR408-3p	miR408	AT3G01480	0.5	Cat_2	55.97	AUGCACUGCCUCUUCCCUGGC	Guide	CYP38	cyclophilin 38	PTHR11071
ath-miR398a-3p	miR398	AT3G15640	0.529	Cat_1	49.79	UGUGUUCUCAGGUCACCCCUU	Guide		Rubredoxin-like superfamily protein	PTHR10122
ath-miR164a	miR164	AT3G12977	0.588	Cat_1	17.48	UGGAGAAGCAGGGCACGUGCA	Guide		NAC (No Apical Meristem) domain transcriptional regulator superfamily protein	PTHR31744
ath-miR156a-5p	miR156/miR157	AT1G53160	0.618	Cat_1	15.41	UGACAGAAGAGAGUGAGCAC	Guide	SPL4	squamosa promoter binding protein-like 4	PTHR31251
ath-miR156a-5p	miR156/miR157	AT5G43270	0.647	Cat_1	35.41	UGACAGAAGAGAGUGAGCAC	Guide	SPL2	squamosa promoter binding protein-like 2	PTHR31251
ath-miR167a-5p	miR167	AT5G58590	0.647	Cat_1	86.84	UGAAGCUGCCAGCAUGAUCUA	Guide	RANBP1	RAN binding protein 1	PTHR23138
ath-miR167a-5p	miR167	AT5G37020	0.647	Cat_1	226.42	UGAAGCUGCCAGCAUGAUCUA	Guide	ARF8	auxin response factor 8	PTHR31384
ath-miR827	miR827	AT1G02860	0.647	Cat_1	39.39	UUAGAUGACCAUCAACAAACU	Guide	NLA	SPX (SYG1/Pho81/XPR1) domain-containing protein	PTHR23041
ath-miR164a	miR164	AT3G15170	0.735	Cat_1	78.92	UGGAGAAGCAGGGCACGUGCA	Guide	CUC1	NAC (No Apical Meristem) domain transcriptional regulator superfamily protein	PTHR31744
ath-miR169a-5p	miR169	AT5G12840	0.735	Cat_1	18.76	CAGCCAAGGAUGACUUGCCGA	Guide	NF-YA1	nuclear factor Y, subunit A1	PTHR12632
ath-miR395a	miR395	AT1G50930	0.735	Cat_1	42.62	CUGAAGUGUUUGGGGGAACUC	Guide		Serine/Threonine-kinase	PTHR33974

ath-miR408-3p	miR408	AT1G72230	0.765	Cat_1	475.39	AUGCACUGCCUCUUCCCUGGC	Guide		Cupredoxin superfamily protein	PTHR33021
ath-miR396a-5p	miR396	AT5G60360	0.794	Cat_2	1010.98	UUCCACAGCUUUCUUGAACUG	Guide	ALP	aleurain-like protease	PTHR12411
ath-miR164a	miR164	AT5G61430	0.824	Cat_1	33.82	UGGAGAAGCAGGGCACGUGCA	Guide	NAC100	NAC domain containing protein 100	PTHR31744
ath-miR167a-5p	miR167	AT1G30330	0.824	Cat_1	53.03	UGAAGCUGCCAGCAUGAUCUA	Guide	ARF6	auxin response factor 6	PTHR31384
ath-miR172a	miR172	AT3G54990	0.882	Cat_1	29.17	AGAAUCUUGAUGAUGCUGCAU	Guide	SMZ	AP2-LIKE ETHYLENE-RESPONSIVE TRANSCRIPTION FACTOR SMZ-RELATED	PTHR32467
ath-miR319a	miR319	AT1G30210	0.912	Cat_1	215.15	UUGGACUGAAGGGAGCUCCCU	Guide	TCP24	TEOSINTE BRANCHED 1, cycloidea, and PCF family 24	PTHR31072
ath-miR395a	miR395	AT5G43780	0.912	Cat_1	39.63	CUGAAGUGUUUGGGGGAACUC	Guide	APS4	Pseudouridine synthase/archaeosine transglycosylase-like family protein	PTHR11055
ath-miR396a-5p	miR396	AT5G43060	0.912	Cat_1	46.02	UUCCACAGCUUUCUUGAACUG	Guide		Granulin repeat cysteine protease family protein	PTHR12411
ath-miR156a-5p	miR156/miR157	AT3G15270	0.941	Cat_1	72.59	UGACAGAAGAGAGUGAGCAC	Guide	SPL5	squamosa promoter binding protein-like 5	PTHR31251
ath-miR396a-5p	miR396	AT3G52910	0.971	Cat_1	34.76	UUCCACAGCUUUCUUGAACUG	Guide	GRF4	growth-regulating factor 4	PTHR31602
ath-miR164a	miR164	AT5G07680	1.029	Cat_1	20.67	UGGAGAAGCAGGGCACGUGCA	Guide	NAC080	NAC domain containing protein 80	PTHR31744
ath-miR396a-5p	miR396	AT3G14110	1.029	Cat_1	59.83	UUCCACAGCUUUCUUGAACUG	Guide	FLU	Tetratricopeptide repeat (TPR)-like superfamily protein	PTHR10098
ath-miR171a-3p	miR170/miR171	AT3G60630	1.147	Cat_1	157.52	UGAUUGAGCCGCGCCAAUAUC	Guide	HAM2	GRAS family transcription factor	PTHR31636
ath-miR397a	miR397	AT3G60250	1.147	Cat_1	192.6	UCAUUGAGUGCAGCGUUGAUG	Guide	CKB3	casein kinase II beta chain 3	PTHR11740
ath-miR398a-3p	miR398	AT2G27530	1.206	Cat_1	190.03	UGUGUUCUCAGGUCACCCCUU	Guide	PGY1	Ribosomal protein L1p/L10e family	PTHR23105
ath-miR168a-5p	miR168	AT3G58030	1.294	Cat_1	13.99	UCGCUUGGUGCAGGUCGGGAA	Guide		RING/U-box superfamily protein	PTHR12313
ath-miR319a	miR319	AT1G53230	1.294	Cat_1	93.72	UUGGACUGAAGGGAGCUCCCU	Guide	TCP3	TEOSINTE BRANCHED 1, cycloidea and PCF transcription factor 3	PTHR31072
ath-miR164a	miR164	AT5G53950	1.324	Cat_1	213.52	UGGAGAAGCAGGGCACGUGCA	Guide	CUC2	NAC (No Apical Meristem) domain transcriptional regulator superfamily protein	PTHR31744
ath-miR156a-5p	miR156/miR157	AT3G57920	1.441	Cat_1	24.75	UGACAGAAGAGAGUGAGCAC	Guide	SPL15	squamosa promoter binding protein-like 15	PTHR31251
ath-miR397a	miR397	AT2G29130	1.471	Cat_1	22.01	UCAUUGAGUGCAGCGUUGAUG	Guide	LAC2	laccase 2	PTHR11709
ath-miR396a-5p	miR396	AT1G10120	1.676	Cat_1	210.45	UUCCACAGCUUUCUUGAACUG	Guide		basic helix-loop-helix (bHLH) DNA-binding superfamily protein	PTHR12565
ath-miR159a	miR159	AT5G06100	1.706	Cat_1	269.64	UUUGGAUUGAAGGGAGCUCUA	Guide	MYB33	myb domain protein 33	PTHR10641
ath-miR162a-3p	miR162	AT1G01040	1.794	Cat_1	116.84	UCGAUAAACCUCUGCAUCCAG	Guide	DCL1	dicer-like 1	PTHR14950
ath-miR408-3p	miR408	AT2G44790	1.912	Cat_1	806.61	AUGCACUGCCUCUUCCCUGGC	Guide	UCC2	uclacyanin 2, a protein precursor that is closely related to precursors of stellacyanins and a blue copper protein	PTHR33021
ath-miR156a-5p	miR156/miR157	AT1G69170	1.971	Cat_1	38.52	UGACAGAAGAGAGUGAGCAC	Guide		Squamosa promoter-binding protein-like (SBP domain) transcription factor family protein	PTHR31251
ath-miR170-3p	miR170/miR171	AT4G00150	1.971	Cat_1	276.2	UGAUUGAGCCGUGUCAAUAUC	Guide	HAM3	GRAS family transcription factor	PTHR31636

ath-miR319a	miR319	AT3G15030	1.971	Cat_1	126.83	UUGGACUGAAGGGAGCUCCCU	Guide	TCP4	TCP family transcription factor 4	PTHR31072
ath-miR398a-3p	miR398	AT2G28190	1.971	Cat_1	674.57	UGUGUUCUCAGGUCACCCCUU	Guide	CSD2	copper/zinc superoxide dismutase 2	PTHR10003
ath-miR169a-5p	miR169	AT5G06510	2.088	Cat_1	83.62	CAGCCAAGGAUGACUUGCCGA	Guide	NF-YA10	nuclear factor Y, subunit A10	PTHR12632
ath-miR169a-5p	miR169	AT3G05690	2.176	Cat_1	154.11	CAGCCAAGGAUGACUUGCCGA	Guide	NF-YA2	nuclear factor Y, subunit A2	PTHR12632
ath-miR172a	miR172	AT5G67180	2.265	Cat_1	207.63	AGAAUCUUGAUGAUGCUGCAU	Guide	TOE3	target of early activation tagged (EAT) 3	PTHR32467
ath-miR395a	miR395	AT3G22890	2.265	Cat_1	350.01	CUGAAGUGUUUGGGGGAACUC	Guide	APS1	ATP sulfurylase 1	PTHR11055
ath-miR172a	miR172	AT2G28550	2.294	Cat_1	320.27	AGAAUCUUGAUGAUGCUGCAU	Guide	RAP2.7	related to AP2.7	PTHR32467
ath-miR168a-5p	miR168	AT1G48410	2.353	Cat_1	106.76	UCGCUUGGUGCAGGUCGGGAA	Guide	AGO1	Argonaute family protein	PTHR22891
ath-miR396a-5p	miR396	AT2G45480	2.353	Cat_1	128.71	UUCCACAGCUUUCUUGAACUG	Guide	GRF9	growth-regulating factor 9	PTHR31602
ath-miR165a-3p	miR165/miR166	AT4G32880	2.382	Cat_1	113.7	UCGGACCAGGCUUCAUCCCCC	Guide	HB-8	homeobox gene 8	PTHR24326
ath-miR169a-5p	miR169	AT1G72830	2.382	Cat_1	139.35	CAGCCAAGGAUGACUUGCCGA	Guide	NF-YA3	nuclear factor Y, subunit A3	PTHR12632
ath-miR156a-5p	miR156/miR157	AT1G27370	2.441	Cat_1	267.76	UGACAGAAGAGAGUGAGCAC	Guide	SPL10	squamosa promoter binding protein-like 10	PTHR31251
ath-miR396a-5p	miR396	AT1G53910	2.559	Cat_1	92.04	UUCCACAGCUUUCUUGAACUG	Guide	RAP2.12	related to AP2 12	PTHR31190
ath-miR156a-5p	miR156/miR157	AT2G33810	2.588	Cat_1	213.52	UGACAGAAGAGAGUGAGCAC	Guide	SPL3	squamosa promoter binding protein-like 3	PTHR31251
ath-miR164a	miR164	AT5G39610	2.647	Cat_1	257.83	UGGAGAAGCAGGGCACGUGCA	Guide	NAC6	NAC domain containing protein 6	PTHR31744
ath-miR393a-5p	miR393	AT3G26810	2.647	Cat_1	106.76	UCCAAAGGGAUCGCAUUGAUCC	Guide	AFB2	auxin signaling F-box 2	PTHR24006
ath-miR159a	miR159	AT3G11440	2.676	Cat_1	269.64	UUUGGAUUGAAGGGAGCUCUA	Guide	MYB65	myb domain protein 65	PTHR10641
ath-miR398a-3p	miR398	AT1G08830	2.794	Cat_1	724.36	UGUGUUCUCAGGUCACCCCUU	Guide	CSD1	copper/zinc superoxide dismutase 1	PTHR10003
ath-miR408-3p	miR408	AT2G02850	2.794	Cat_1	389.9	AUGCACUGCCUCUUCCCUGGC	Guide	ARPN	plantacyanin	PTHR33021
ath-miR170-3p	miR170/miR171	AT2G45160	2.853	Cat_1	4348.85	UGAUUGAGCCGUGUCAAUAUC	Guide	HAM1	GRAS family transcription factor	PTHR31636
ath-miR156a-5p	miR156/miR157	AT5G50570	2.882	Cat_1	533.79	UGACAGAAGAGAGUGAGCAC	Guide		Squamosa promoter-binding protein-like (SBP domain) transcription factor family protein	PTHR31251
ath-miR156g	miR156/miR157	AT5G50670	2.882	Cat_1	533.79	CGACAGAAGAGAGUGAGCAC	Guide		Squamosa promoter-binding protein-like (SBP domain) transcription factor family protein	PTHR31251
ath-miR393a-5p	miR393	AT4G03190	2.882	Cat_1	100.53	UCCAAAGGGAUCGCAUUGAUCC	Guide	GRH1	GRR1-like protein 1	PTHR24006
ath-miR393a-5p	miR393	AT3G62980	3	Cat_1	247.69	UCCAAAGGGAUCGCAUUGAUCC	Guide	TIR1	F-box/RNI-like superfamily protein	PTHR24006
ath-miR396a-5p	miR396	AT2G36400	3.176	Cat_1	223.31	UUCCACAGCUUUCUUGAACUG	Guide	GRF3	growth-regulating factor 3	PTHR31602
ath-miR398a-3p	miR398	AT3G27200	3.176	Cat_1	453.78	UGUGUUCUCAGGUCACCCCUU	Guide		Cupredoxin superfamily protein	PTHR33021
ath-miR169a-5p	miR169	AT1G17590	3.265	Cat_1	201.11	CAGCCAAGGAUGACUUGCCGA	Guide	NF-YA8	nuclear factor Y, subunit A8	PTHR12632
ath-miR160a-5p	miR160	AT2G28350	3.294	Cat_1	523.68	UGCCUGGCUCCCUGUAUGCCA	Guide	ARF10	auxin response factor 10	PTHR31384
ath-miR398a-3p	miR398	AT1G12520	3.324	Cat_1	2490.65	UGUGUUCUCAGGUCACCCCUU	Guide	CCS	copper chaperone for SOD1	PTHR10003
ath-miR156a-5p	miR156/miR157	AT1G27360	3.353	Cat_1	236.76	UGACAGAAGAGAGUGAGCAC	Guide	SPL11	squamosa promoter-like 11	PTHR31251

ath-miR172e-3p	miR172	AT4G36920	3.412	Cat_1	356.07	GGAAUCUUGAUGAUGCUGCAU	Guide	AP2	Integrase-type DNA-binding superfamily protein (FLORAL HOMEOTIC PROTEIN APETALA 2)	PTHR32467
ath-miR396a-5p	miR396	AT4G24150	3.529	Cat_1	1086.06	UUCCACAGCUUUCUUGAACUG	Guide	GRF8	growth-regulating factor 8	PTHR31602
ath-miR164a	miR164	AT1G56010	3.559	Cat_1	624.63	UGGAGAAGCAGGGCACGUGCA	Guide	NAC1	NAC domain containing protein 1	PTHR31744
ath-miR165a-3p	miR165/ miR166	AT5G60690	3.559	Cat_1	295.07	UCGGACCAGGCUUCAUCCCCC	Guide	REV	Homeobox-leucine zipper family protein / lipid-binding START domain-containing protein	PTHR24326
ath-miR165a-3p	miR165/ miR166	AT1G52150	3.618	Cat_1	424.21	UCGGACCAGGCUUCAUCCCCC	Guide	ATHB-15	Homeobox-leucine zipper family protein / lipid-binding START domain-containing protein	PTHR24326
ath-miR403-3p	miR403	AT1G31280	3.706	Cat_1	636.99	UUAGAUUCACGCACAAACUCG	Guide	AGO2	Argonaute family protein	PTHR22891
ath-miR396a-5p	miR396	AT2G22840	3.765	Cat_1	673.63	UUCCACAGCUUUCUUGAACUG	Guide	GRF1	growth-regulating factor 1	PTHR31602
ath-miR165a-3p	miR165/ miR166	AT1G30490	3.882	Cat_1	1140.94	UCGGACCAGGCUUCAUCCCCC	Guide	PHV	Homeobox-leucine zipper family protein / lipid-binding START domain-containing protein	PTHR24326
ath-miR396a-5p	miR396	AT4G37740	3.971	Cat_1	308.16	UUCCACAGCUUUCUUGAACUG	Guide	GRF2	growth-regulating factor 2	PTHR31602
ath-miR165a-3p	miR165/ miR166	AT2G34710	4	Cat_1	1223.06	UCGGACCAGGCUUCAUCCCCC	Guide	PHB	Homeobox-leucine zipper family protein / lipid-binding START domain-containing protein	PTHR24326
ath-miR172a	miR172	AT5G60120	4	Cat_1	709.33	AGAAUCUUGAUGAUGCUGCAU	Guide	TOE2	target of early activation tagged (EAT) 2	PTHR32467
ath-miR160a-5p	miR160	AT1G77850	4.118	Cat_1	619.8	UGCCUGGCUCCCUGUAUGCCA	Guide	ARF17	auxin response factor 17	PTHR31384
ath-miR393a-5p	miR393	AT1G12820	4.118	Cat_1	731.71	UCCAAAGGGAUCGCAUUGAUCC	Guide	AFB3	auxin signaling F-box 3	PTHR24006
ath-miR160a-5p	miR160	AT4G30080	4.294	Cat_1	725.3	UGCCUGGCUCCCUGUAUGCCA	Guide	ARF16	auxin response factor 16	PTHR31384
ath-miR2111a-3p	miR2111	No HE Targets	NA	NA	NA	GUCCUCGGGAUGCGGAUUACC	Guide	NA	NA	NA
ath-miR391-3p	miR391	No HE Targets	NA	NA	NA	ACGGUAUCUCUCCUACGUAGC	Guide	NA	NA	NA
ath-miR828	miR828	No HE Targets	NA	NA	NA	UCUUGCUUAAAUGAGUAUUCCA	Guide	NA	NA	NA
ath-miR845a	miR845	No HE Targets	NA	NA	NA	CGGCUCUGAUACCAAUUGAUG	Guide	NA	NA	NA

Table S5. All HE targets of *Brassicaceae* specific miRNAs and their Category Scores

Targets which are part of or related to targets of the VAT are highlighted in blue

Targets of miRNAs previously reported in literature, or are related to these targets, are highlighted in purple (this includes miR400, miR472, miR858)

*The highest cleavage tag abundance found for this gene across all degradome libraries

miRNA	miRNA family	Gene ID	Conservation	Cat score	Maximum Category	Cleavage tag abundance*	miRNA sequence	Gene symbol	Gene brief description	Validation method	Reference
ath-miR837-5p	miR837	AT1G01160	Brassicaceae	0.118	Cat_2	6.92	AUCAGUUUCUUGUUCGUUUCA	GIF2	GRF1-interacting factor 2		
ath-miR835-5p	miR835	AT1G51805	Brassicaceae	0.118	Cat_2	23.27	UUCUUGCAUAUGUUCUUUAUC		Leucine-rich repeat protein kinase family protein		
ath-miR842	miR842	AT1G62750	Brassicaceae	0.118	Cat_2	17.31	UCAUGGUCAGAUCCGUCAUCC	SCO1	Translation elongation factor EFG/EF2 protein		
ath-miR400	miR400	AT1G63080	Brassicaceae	0.118	Cat_2	18.4	UAUGAGAGUAUUAUAAGUCAC		Pentatricopeptide repeat (PPR) superfamily protein		
ath-miR400	miR400	AT1G63130	Brassicaceae	0.118	Cat_2	18.4	UAUGAGAGUAUUAUAAGUCAC		Tetratricopeptide repeat (TPR)-like superfamily protein		
ath-miR829-3p.2	miR829	AT1G64170	Brassicaceae	0.118	Cat_2	6.6	CAAAUUAAAGCUUCAAGGUAG	CHX16	cation/H+ exchanger 16		
ath-miR829-5p	miR829	AT2G25520	Brassicaceae	0.118	Cat_2	27.98	ACUUUGAAGCUUUGAUUUGAA		Drug/metabolite transporter superfamily protein		
ath-miR4243	miR4243	AT3G02830	Brassicaceae	0.118	Cat_2	15.12	UUGAAAUUGUAGAUUUCGUAC	ZFN1	zinc finger protein 1		
ath-miR158a-3p	miR158	AT4G10770	Brassicaceae	0.118	Cat_2	20.58	UCCCAAAUGUAGACAAAGCA	OPT7	oligopeptide transporter 7		
ath-miR173-3p	miR173	AT5G26830	Brassicaceae	0.118	Cat_2	8.84	UGAUUCUCUGUGUAAGCGAAA		Threonyl-tRNA synthetase		
ath-miR838	miR838	AT1G51630	Brassicaceae	0.147	Cat_2	19.73	UUUUCUUCUACUUCUUGCACA		O-fucosyltransferase family protein		
ath-miR774b-3p	miR774	AT2G39795	Brassicaceae	0.147	Cat_2	74.74	CAUCCAUAUUUUCAUCUCGAA		Mitochondrial glycoprotein family protein		
ath-miR8171	miR8171	AT3G52840	Brassicaceae	0.147	Cat_2	9.2	AUAGGUGGGCCAGUGGUAGGA	BGAL2	beta-galactosidase 2		
ath-miR857	miR857	AT5G36880	Brassicaceae	0.147	Cat_2	10.83	UUUUGUAUGUUGAAGGUGUAU	ACS	acetyl-CoA synthetase		
ath-miR831-3p	miR831	AT3G09630	Brassicaceae	0.176	Cat_2	18.58	UGAUCUCUUCGUACUCUUCUUG		Ribosomal protein L4/L1 family		
ath-miR831-3p	miR831	AT3G11120	Brassicaceae	0.176	Cat_2	21.45	UGAUCUCUUCGUACUCUUCUUG		Ribosomal protein L41 family		

ath-miR869.2	miR869	AT3G14310	Brassicaceae	0.176	Cat_2	13.18	UCUGGUGUUGAGAUAGUUGAC	PME3	pectin methylesterase 3		
ath-miR774b-3p	miR774	AT3G16770	Brassicaceae	0.176	Cat_2	22.52	CAUCCAUAUUUUCAUCUCGAA	EBP	ethylene-responsive element binding protein		
ath-miR833b	miR833	AT1G45145	Brassicaceae	0.206	Cat_2	13.12	UGUUUGUUGACAUCGGUCUAG	TRX5	thioredoxin H-type 5		
ath-miR161.2	miR161	AT2G41720	Brassicaceae	0.206	Cat_2	25.97	UCAAUGCAUUGAAAGUGACUA	EMB2654	Tetratricopeptide repeat (TPR)-like superfamily protein		
ath-miR472-5p	miR472	AT4G33010	Brassicaceae	0.206	Cat_2	34.24	AUGGUCGAAGUAGGCAAAAUC	GLDP1	glycine decarboxylase P-protein 1		
ath-miR852	miR852	AT1G79920	Brassicaceae	0.235	Cat_2	25.33	AAGAUAAGCGCCUUAGUUCUG		Heat shock protein 70 (Hsp 70) family protein		
ath-miR838	miR838	AT3G63240	Brassicaceae	0.235	Cat_2	9.4	UUUUCUUCUACUUCUUGCACA		DNAse I-like superfamily protein		
ath-miR868-5p	miR868	AT1G10150	Brassicaceae	0.265	Cat_2	13.01	UCAUGUCGUAAUAGUAGUCAC		Carbohydrate-binding protein		
ath-miR831-5p	miR831	AT1G21460	Brassicaceae	0.265	Cat_2	23.01	AGAAGCGUACAAGGAGAUGAGG		Nodulin MtN3 family protein		
ath-miR402	miR402	AT1G22610	Brassicaceae	0.265	Cat_1	18.88	UUCGAGGCCUAUUAAACCUCUG		C2 calcium/lipid-binding plant phosphoribosyltransferase family protein		
ath-miR4245	miR4245	AT1G27370	Brassicaceae	0.265	Cat_1	213.52	ACAAAGUUUUAUACUGACAAU	SPL10	squamosa promoter binding protein-like 10		
ath-miR161.2	miR161	AT1G64780	Brassicaceae	0.265	Cat_2	30.81	UCAAUGCAUUGAAAGUGACUA	AMT1;2	ammonium transporter 1;2		
ath-miR158a-5p	miR158	AT3G22630	Brassicaceae	0.265	Cat_1	267.39	CUUUGUCUACAAUUUUGGAAA	PBD1	20S proteasome beta subunit D1		
ath-miR161.2	miR161	AT1G63400	Brassicaceae	0.294	Cat_1	6.29	UCAAUGCAUUGAAAGUGACUA		Pentatricopeptide repeat (PPR) superfamily protein		
ath-miR840-3p	miR840	AT4G30190	Brassicaceae	0.324	Cat_2	36.07	UUGUUUAGGUCCCUUAGUUUC	HA2	H(+)-ATPase 2		
ath-miR860	miR860	AT1G71840	Brassicaceae	0.353	Cat_1	69.05	UCAAUAGAUUGGACUAUGUAU		transducin family protein / WD-40 repeat family protein		
ath-miR834	miR834	AT3G14310	Brassicaceae	0.353	Cat_2	28.19	UGGUAGCAGUAGCGGUGGUAA	PME3	pectin methylesterase 3		
ath-miR857	miR857	AT3G09220	Brassicaceae	0.412	Cat_1	10.39	UUUUGUAUGUUGAAGGUGUAU	LAC7	laccase 7		

ath-miR173-5p	miR173	AT4G39090	Brassicaceae	0.441	Cat_1	87.44	UUCGCUUGCAGAGAGAAAUCAC	RD19	Papain family cysteine protease		
ath-miR858b	miR858	AT1G22640	Brassicaceae	0.471	Cat_2	37.99	UUCGUUGUCUGUUCGACCUUG	MYB3	myb domain protein 3	Related to MYB12	
ath-miR3434-5p	miR3434	AT1G37130	Brassicaceae	0.471	Cat_1	34.24	ACUUGGCUGAUUCUAUUAUU	NIA2	nitrate reductase 2		German et al., 2008; Boccara et al., 2014
ath-miR173-3p	miR173	AT5G58590	Brassicaceae	0.471	Cat_1	16.94	UGAUUCUCUGUGUAAGCGAAA	RANBP1	RAN binding protein 1		
ath-miR858b	miR858	AT4G26930	Brassicaceae	0.559	Cat_1	17.9	UUCGUUGUCUGUUCGACCUUG	MYB97	myb domain protein 97	Related to MYB12	
ath-miR161.1	miR161	AT1G62910	Brassicaceae	0.588	Cat_2	41.57	UGAAAGUGACUACAUCGGGGU		Pentatricopeptide repeat (PPR) superfamily protein		
ath-miR161.1	miR161	AT1G62914	Brassicaceae	0.588	Cat_2	41.57	UGAAAGUGACUACAUCGGGGU		pentatricopeptide (PPR) repeat-containing protein		Park et al., 2014
ath-miR161.1	miR161	AT1G62930	Brassicaceae	0.588	Cat_2	41.57	UGAAAGUGACUACAUCGGGGU		Tetratricopeptide repeat (TPR)-like superfamily protein		
ath-miR161.1	miR161	AT1G63130	Brassicaceae	0.588	Cat_2	41.57	UGAAAGUGACUACAUCGGGGU		Tetratricopeptide repeat (TPR)-like superfamily protein		
ath-miR163	miR163	AT3G44860	Brassicaceae	0.588	Cat_1	43.86	UUGAAGAGGACUUGGAACUUCGAU	FAMT	farnesoic acid carboxyl-O-methyltransferase		
ath-miR858a	miR858	AT2G47460	Brassicaceae	0.618	Cat_1	38.39	UUUCGUUGUCUGUUCGACCUU	MYB12	myb domain protein 12	degradome, correlation of miRNA/target mRNA levels, miRNA OE, MIMIC OE	German et al., 2008; Sharma et al., 2014
ath-miR868-3p	miR868	AT1G18270	Brassicaceae	0.676	Cat_1	14.44	CUUCUUAAGUGCUGAUAAUGC		ketose-bisphosphate aldolase class-II family protein		
ath-miR831-3p	miR831	AT3G08520	Brassicaceae	0.676	Cat_2	90.29	UGAUCUCUUCGUACUCUUCUUG		Ribosomal protein L41 family		
ath-miR400	miR400	AT1G62720	Brassicaceae	0.735	Cat_1	18.73	UAUGAGAGUAUUAUAAGUCAC		Pentatricopeptide repeat (PPR-like) superfamily protein	5' RACE, miRNA OE	
ath-miR831-3p	miR831	AT3G56020	Brassicaceae	0.735	Cat_2	90.29	UGAUCUCUUCGUACUCUUCUUG		Ribosomal protein L41 family		

ath-miR4221	miR4221	AT1G20500	Brassicaceae	1.059	Cat_1	42.43	UUUUCCUCUGUUGAAUUCUUGC		AMP-dependent synthetase and ligase family protein		
ath-miR161.1	miR161	AT1G64583	Brassicaceae	1.059	Cat_1	10.39	UGAAAGUGACUACAUCGGGGU		Tetratricopeptide repeat (TPR)-like superfamily protein		
ath-miR163	miR163	AT1G66700	Brassicaceae	1.206	Cat_1	28.89	UUGAAGAGGACUUGGAACUUCGAU	PXMT1	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein		
ath-miR472-3p	miR472	AT5G43740	Brassicaceae	1.529	Cat_1	40.53	UUUUUCCUACUCCGCCCAUACC		Disease resistance protein (CC-NBS-LRR class) family	Degradome, miRNA OE, *rdr6* mutant	
ath-miR161.1	miR161	AT1G06580	Brassicaceae	1.794	Cat_1	33.49	UGAAAGUGACUACAUCGGGGU		Pentatricopeptide repeat (PPR) superfamily protein		
ath-miR823	miR823	AT1G69770	Brassicaceae	2.294	Cat_1	90.15	UGGGUGGUGAUCAUAUAAGAU	CMT3	chromomethylase 3		
ath-miR824-5p	miR824	AT3G57230	Brassicaceae	3.471	Cat_1	858.71	UAGACCAUUUGUGAGAAGGGA	AGL16	AGAMOUS-like 16		
ath-miR161.2	miR161	AT5G41170	Brassicaceae	4.118	Cat_1	280.52	UCAAUGCAUUGAAAGUGACUA		Pentatricopeptide repeat (PPR-like) superfamily protein		
ath-miR5654-3p	miR5654	No HE Targets	Brassicaceae	NA	NA	NA	UGGAAGAUGCUUUGGGAUUUAUU	NA	NA		
ath-miR774a	miR774	No HE Targets	Brassicaceae	NA	NA	NA	UUGGUUACCCAUAUGGCCAUC	NA	NA		
ath-miR781a	miR781	No HE Targets	Brassicaceae	NA	NA	NA	UUAGAGUUUUCUGGAUACUUA	NA	NA		
ath-miR822-3p	miR822	No HE Targets	Brassicaceae	NA	NA	NA	UGUGCAAAUGCUUUCUACAGG	NA	NA		
ath-miR824-3p	miR824	No HE Targets	Brassicaceae	NA	NA	NA	CCUUCUCAUCGAUGGUCUAGA	NA	NA		
ath-miR1887	miR1887	No HE Targets	Brassicaceae	NA	NA	NA	UACUAAGUAGAGUCUAAGAGA	NA	NA		
ath-miR2112-3p	miR2112	No HE Targets	Brassicaceae	NA	NA	NA	CUUUAUAUCCGCAUUUGCGCA	NA	NA		
ath-miR3434-3p	miR3434	No HE Targets	Brassicaceae	NA	NA	NA	UCAGAGUAUCAGCCAUGUGA	NA	NA		
ath-miR3440b-3p	miR3440	No HE Targets	Brassicaceae	NA	NA	NA	UGGAUUGGUCAAGGGAAGCGU	NA	NA		

ath-miR4227	miR4227	No HE Targets	Brassicaceae	NA	NA	NA	UCACUGGUACCAAUCAUUCCA	NA	NA		
ath-miR4228-3p	miR4228	No HE Targets	Brassicaceae	NA	NA	NA	UCGGAUGCGAAACGGUGGUGU	NA	NA		
ath-miR4239	miR4239	No HE Targets	Brassicaceae	NA	NA	NA	UUUGUUAUUUUCGCAUGCUCC	NA	NA		
ath-miR4240	miR4240	No HE Targets	Brassicaceae	NA	NA	NA	UGACUAGACCCGUAACAUUAC	NA	NA		
ath-miR844-3p	miR844	No HE Targets	Brassicaceae	NA	NA	NA	UUAUAAGCCAUCUUACUAGUU	NA	NA		
ath-miR846-3p	miR846	No HE Targets	Brassicaceae	NA	NA	NA	UUGAAUUGAAGUGCUUGAAUU	NA	NA		
ath-miR847	miR847	No HE Targets	Brassicaceae	NA	NA	NA	UCACUCCUCUUCUUCUUGAUG	NA	NA		
ath-miR848	miR848	No HE Targets	Brassicaceae	NA	NA	NA	UGACAUGGGACUGCCUAAGCUA	NA	NA		
ath-miR851-3p	miR851	No HE Targets	Brassicaceae	NA	NA	NA	UGGGUGGCAAACAAAGACGAC	NA	NA		
ath-miR825	miR825	No HE Targets	Brassicaceae	NA	NA	NA	UUCUCAAGAAGGUGCAUGAAC	NA	NA		
ath-miR829-3p.1	miR829	No HE Targets	Brassicaceae	NA	NA	NA	AGCUCUGAUACCAAAUGAUGGAAU	NA	NA		
ath-miR833a-3p	miR833	No HE Targets	Brassicaceae	NA	NA	NA	UAGACCGAUGUCAACAAACAAG	NA	NA		
ath-miR835-3p	miR835	No HE Targets	Brassicaceae	NA	NA	NA	UGGAGAAGAUACGCAAGAAAG	NA	NA		
ath-miR837-3p	miR837	No HE Targets	Brassicaceae	NA	NA	NA	AAACGAACAAAAAACUGAUGG	NA	NA		
ath-miR839-5p	miR839	No HE Targets	Brassicaceae	NA	NA	NA	UACCAACCUUUCAUCGUUCCC	NA	NA		
ath-miR840-5p	miR840	No HE Targets	Brassicaceae	NA	NA	NA	ACACUGAAGGACCUAAACUAAC	NA	NA		
ath-miR841a-3p	miR841	No HE Targets	Brassicaceae	NA	NA	NA	AUUUCUAGUGGGUCGUAUUCA	NA	NA		
ath-miR861-3p	miR861	No HE Targets	Brassicaceae	NA	NA	NA	GAUGGAUAUGUCUUCAAGGAC	NA	NA		
ath-miR853	miR853	No HE Targets	Brassicaceae	NA	NA	NA	UCCCCUCUUUAGCUUGGAGAAG	NA	NA		
ath-miR856	miR856	No HE Targets	Brassicaceae	NA	NA	NA	UAAUCCUACCAAUAACUUCAGC	NA	NA		

ath-miR859	miR859	No HE Targets	Brassicaceae	NA	NA	NA	UCUCUCUGUUGUGAAGUCAAA	NA	NA		
ath-miR862-3p	miR862	No HE Targets	Brassicaceae	NA	NA	NA	AUAUGCUGGAUCUACUUGAAG	NA	NA		
ath-miR869.1	miR869	No HE Targets	Brassicaceae	NA	NA	NA	AUUGGUUCAAUUCUGGUGUUG	NA	NA		

Table S6. All HE targets of *A. thaliana* specific miRNAs and their Category Scores

Orange indicates miRNAs with repetitive sequences

*The highest cleavage tag abundance found for this gene across all degradome libraries

miRNA	miRNA family	Gene ID	Conservation	Cat score	Maximum Category	Cleavage tag abundance*	miRNA sequence	Gene symbol	Gene brief description
ath-miR775	miR775	AT1G12240	A_thaliana_only	0.118	Cat_2	17.73	UUCGAUGUCUAGCAGUGCCA	ATBETAFRUCT4	Glycosyl hydrolases family 32 protein
ath-miR5021	miR5021	AT1G70900	A_thaliana_only	0.118	Cat_2	24.72	UGAGAAGAAGAAGAAGAAAA		(1 of 2) PTHR35100:SF1 - F15H11.13 PROTEIN
ath-miR5631	miR5631	AT2G20260	A_thaliana_only	0.118	Cat_2	15.26	UGGCAGGAAAGACAUAAUUUU	PSAE-2	photosystem I subunit E-2
ath-miR863-3p	miR863	AT3G16470	A_thaliana_only	0.118	Cat_2	24.79	UUGAGAGCAACAAGACAUAAU	JR1	Mannose-binding lectin superfamily protein
ath-miR414	miR414	AT3G19910	A_thaliana_only	0.118	Cat_2	12.1	UCAUCUUCAUCAUCAUCGUCA		RING/U-box superfamily protein
ath-miR5020a	miR5020	AT3G55960	A_thaliana_only	0.118	Cat_2	44.4	UGGAAGAAGGUGAGACUUGCA		Haloacid dehalogenase-like hydrolase (HAD) superfamily protein
ath-miR5644	miR5644	AT4G02510	A_thaliana_only	0.118	Cat_2	19.14	GUGGGUUGCGGAUAACGGUA	TOC159	translocon at the outer envelope membrane of chloroplasts 159
ath-miR414	miR414	AT4G12610	A_thaliana_only	0.118	Cat_2	26.55	UCAUCUUCAUCAUCAUCGUCA	RAP74	transcription activators;DNA binding;RNA polymerase II transcription factors;catalytics;transcription initiation factors
ath-miR8170-5p	miR8170	AT4G30190	A_thaliana_only	0.118	Cat_2	12.64	AUAGCAAAUCGAUAAGCAAUG	HA2	H(+)-ATPase 2
ath-miR826a	miR826	AT5G24930	A_thaliana_only	0.118	Cat_2	13.43	UAGUCCGGUUUUGGAUACGUG	COL4	CONSTANS-like 4
ath-miR826b	miR826	AT5G38030	A_thaliana_only	0.118	Cat_2	7.65	UGGUUUUGGACACGUGAAAAU		MATE efflux family protein
ath-miR8167a	miR8167	AT5G43460	A_thaliana_only	0.118	Cat_2	12.81	AGAUGUGGAGAUCGUGGGGAUG		HR-like lesion-inducing protein-related
ath-miR780.1	miR780	AT5G61820	A_thaliana_only	0.118	Cat_2	15.12	UCUAGCAGCUGUUGAGCAGGU		(1 of 1) PF07712 - Stress up-regulated Nod 19 (SURNod19)
ath-miR406	miR406	AT1G11130	A_thaliana_only	0.147	Cat_2	12.02	UAGAAUGCUAUUGUAAUCCAG	SUB	Leucine-rich repeat protein kinase family protein
ath-miR5636	miR5636	AT2G21330	A_thaliana_only	0.147	Cat_2	20.2	CGUAGUUGCAGAGCUUGACGG	FBA1	fructose-bisphosphate aldolase 1
ath-miR2937	miR2937	AT3G14230	A_thaliana_only	0.147	Cat_2	13.2	AUAAGAGCUGUUGAAGGAGUC	RAP2.2	related to AP2 2
ath-miR8183	miR8183	AT3G16470	A_thaliana_only	0.147	Cat_2	14.76	UUUAGUUGACGGAAUUGUGGC	JR1	Mannose-binding lectin superfamily protein
ath-miR5021	miR5021	AT3G26570	A_thaliana_only	0.147	Cat_2	27.99	UGAGAAGAAGAAGAAGAAAA	PHT2;1	phosphate transporter 2;1
ath-miR10515	miR10515	AT3G45600	A_thaliana_only	0.147	Cat_2	10.41	ACCCCGAUGGUUAUCCUCACC	TET3	tetraspanin3
ath-miR5658	miR5658	AT4G23680	A_thaliana_only	0.147	Cat_2	18.04	AUGAUGAUGAUGAUGAUGAAA		Polyketide cyclase/dehydrase and lipid transport superfamily protein
ath-miR5658	miR5658	AT5G14740	A_thaliana_only	0.147	Cat_2	16.73	AUGAUGAUGAUGAUGAUGAAA	CA2	carbonic anhydrase 2
ath-miR414	miR414	AT5G55920	A_thaliana_only	0.147	Cat_2	20.28	UCAUCUUCAUCAUCAUCGUCA	OLI2	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
ath-miR2933a	miR2933	AT1G23490	A_thaliana_only	0.176	Cat_2	8.95	GAAAUCGGAGAGGAAAUUCGCC	ARF1	ADP-ribosylation factor 1

ath-miR5630a	miR5630	AT1G33120	A_thaliana_only	0.176	Cat_2	18.81	GCUAAGAGCGGUUCUGAUGGA		Ribosomal protein L6 family
ath-miR5630a	miR5630	AT1G33140	A_thaliana_only	0.176	Cat_2	18.81	GCUAAGAGCGGUUCUGAUGGA	PGY2	Ribosomal protein L6 family
ath-miR5650	miR5650	AT1G68560	A_thaliana_only	0.176	Cat_2	15.58	UUGUUUUGGAUCUUAGAUACA	XYL1	alpha-xylosidase 1
ath-miR416	miR416	AT1G79350	A_thaliana_only	0.176	Cat_2	29.5	GGUUCGUACGUACACUGUUCA	EMB1135	RING/FYVE/PHD zinc finger superfamily protein
ath-miR5021	miR5021	AT2G46340	A_thaliana_only	0.176	Cat_2	55.97	UGAGAAGAAGAAGAAGAAAA	SPA1	SPA (suppressor of phyA-105) protein family
ath-miR5658	miR5658	AT3G01500	A_thaliana_only	0.176	Cat_2	30.54	AUGAUGAUGAUGAUGAUGAAA	CA1	carbonic anhydrase 1
ath-miR866-3p	miR866	AT3G59780	A_thaliana_only	0.176	Cat_2	34.21	ACAAAAUCCGUCUUUGAAGA		Rhodanese/Cell cycle control phosphatase superfamily protein
ath-miR5021	miR5021	AT3G60680	A_thaliana_only	0.176	Cat_2	55.97	UGAGAAGAAGAAGAAGAAAA		Plant protein of unknown function (DUF641)
ath-miR826b	miR826	AT4G02510	A_thaliana_only	0.176	Cat_2	18.34	UGGUUUUGGACACGUGAAAAU	TOC159	translocon at the outer envelope membrane of chloroplasts 159
ath-miR5651	miR5651	AT4G31850	A_thaliana_only	0.176	Cat_2	10.13	UUGUGCGGUUCAAAUAGUAAC	PGR3	proton gradient regulation 3
ath-miR864-3p	miR864	AT5G52640	A_thaliana_only	0.176	Cat_2	20.26	UAAAGUCAAUAAUACCUUGAAG	HSP90.1	heat shock protein 90.1
ath-miR414	miR414	AT1G63980	A_thaliana_only	0.206	Cat_2	19.53	UCAUCUUCAUCAUCAUCGUCA		D111/G-patch domain-containing protein
ath-miR5024-5p	miR5024	AT2G47115	A_thaliana_only	0.206	Cat_2	21.67	AUGACAAGGCCAAGAUAUAACA		
ath-miR414	miR414	AT3G49140	A_thaliana_only	0.206	Cat_2	27.98	UCAUCUUCAUCAUCAUCGUCA		Pentatricopeptide repeat (PPR) superfamily protein
ath-miR843	miR843	AT4G25640	A_thaliana_only	0.206	Cat_2	17.46	UUUAGGUCGAGCUUCAUUGGA	DTX35	detoxifying efflux carrier 35
ath-miR5657	miR5657	AT5G27660	A_thaliana_only	0.206	Cat_2	25.27	UGGACAAGGUUAGAUUUGGUG		Trypsin family protein with PDZ domain
ath-miR773a	miR773	AT1G30530	A_thaliana_only	0.235	Cat_2	19.5	UUUGCUUCCAGCUUUUGUCUC	UGT78D1	UDP-glucosyl transferase 78D1
ath-miR5658	miR5658	AT5G03060	A_thaliana_only	0.235	Cat_2	17.46	AUGAUGAUGAUGAUGAUGAAA		
ath-miR414	miR414	AT2G11910	A_thaliana_only	0.265	Cat_2	35.8	UCAUCUUCAUCAUCAUCGUCA		
ath-miR5653	miR5653	AT2G39700	A_thaliana_only	0.265	Cat_2	19.22	UGGGUUGAGUUGAGUUGAGUUGGC	EXPA4	expansin A4
ath-miR5629	miR5629	AT5G17770	A_thaliana_only	0.265	Cat_2	17.15	UUAGGGUAGUUAACGGAAGUUA	CBR	NADH:cytochrome B5 reductase 1
ath-miR865-3p	miR865	AT1G07320	A_thaliana_only	0.294	Cat_2	19.69	UUUUUCCUCAAAUUUAUCCAA	RPL4	ribosomal protein L4
ath-miR415	miR415	AT1G15690	A_thaliana_only	0.294	Cat_1	27.99	AACAGAGCAGAAACAGAACAU	AVP1	Inorganic H pyrophosphatase family protein
ath-miR773b-3p	miR773	AT5G04140	A_thaliana_only	0.294	Cat_2	29.59	UUUGAUUCCAGCUUUUGUCUC	GLU1	glutamate synthase 1
ath-miR5661	miR5661	AT5G05200	A_thaliana_only	0.294	Cat_2	43.29	AGAGGUACAUCAUGUAGUCUG		Protein kinase superfamily protein
ath-miR414	miR414	AT3G28850	A_thaliana_only	0.324	Cat_1	39.28	UCAUCUUCAUCAUCAUCGUCA		Glutaredoxin family protein
ath-miR865-5p	miR865	AT1G80780	A_thaliana_only	0.353	Cat_1	34.28	AUGAAUUUGGAUCUAAUUGAG		Polynucleotidyl transferase, ribonuclease H-like superfamily protein
ath-miR8181	miR8181	AT1G54740	A_thaliana_only	0.382	Cat_1	28.57	UGGGGGUGGGGGGGUGACAG		Protein of unknown function (DUF3049)
ath-miR414	miR414	AT2G27170	A_thaliana_only	0.382	Cat_1	148.72	UCAUCUUCAUCAUCAUCGUCA	TTN7	Structural maintenance of chromosomes (SMC) family protein
ath-miR1886.1	miR1886	AT2G47940	A_thaliana_only	0.382	Cat_1	145.62	UGAGAGAAGUGAGAUGAAAUC		DEGP protease 2
ath-miR832-3p	miR832	AT5G13630	A_thaliana_only	0.382	Cat_2	62.31	UUGAUUCCCAAUCCAAGCAAG	GUN5	magnesium-chelatase subunit chlH, chloroplast, putative / Mg-protoporphyrin IX chelatase, putative (CHLH)

ath-miR414	miR414	AT1G50410	A_thaliana_only	0.412	Cat_1	702.06	UCAUCUUCAUCAUCAUCGUCA		SNF2 domain-containing protein / helicase domain-containing protein / zinc finger protein-related
ath-miR5022	miR5022	AT4G34980	A_thaliana_only	0.412	Cat_2	32.79	GUCAUGGGGUAUGAUCGAAUG	SLP2	subtilisin-like serine protease 2
ath-miR850	miR850	AT2G04842	A_thaliana_only	0.471	Cat_2	24.95	UAAGAUCCGGACUACAACAAAG	EMB2761	threonyl-tRNA synthetase, putative / threonine--tRNA ligase, putative
ath-miR5024-3p	miR5024	AT5G03240	A_thaliana_only	0.471	Cat_2	39.03	CCGUAUCUUGGCCUUGUCAUU	UBQ3	polyubiquitin 3
ath-miR5024-3p	miR5024	AT5G20620	A_thaliana_only	0.471	Cat_2	39.03	CCGUAUCUUGGCCUUGUCAUU	UBQ4	ubiquitin 4
ath-miR826a	miR826	AT1G09730	A_thaliana_only	0.5	Cat_1	15.41	UAGUCCGGUUUUGGAUACGUG		Cysteine proteinases superfamily protein
ath-miR5024-3p	miR5024	AT3G57290	A_thaliana_only	0.5	Cat_2	40.53	CCGUAUCUUGGCCUUGUCAUU	EIF3E	eukaryotic translation initiation factor 3E
ath-miR8180	miR8180	AT4G29350	A_thaliana_only	0.5	Cat_1	10.33	UGCGGUGCGGGAGAAGUGC	PFN2	profilin 2
ath-miR5650	miR5650	AT5G20620	A_thaliana_only	0.5	Cat_2	97.6	UUGUUUUGGAUCUUAGAUACA	UBQ4	ubiquitin 4
ath-miR5658	miR5658	AT4G20070	A_thaliana_only	0.618	Cat_1	69.36	AUGAUGAUGAUGAUGAUGAAA	AAH	allantoate amidohydrolase
ath-miR5650	miR5650	AT5G03240	A_thaliana_only	0.618	Cat_2	201.54	UUGUUUUGGAUCUUAGAUACA	UBQ3	polyubiquitin 3
ath-miR2934-3p	miR2934	AT5G03650	A_thaliana_only	0.676	Cat_1	116.02	CAUCCAAGGUGUUUGUAGAAA	SBE2.2	starch branching enzyme 2.2
ath-miR8183	miR8183	AT5G04220	A_thaliana_only	0.676	Cat_1	9.61	UUUAGUUGACGGAAUUGUGGC	SYTC	Calcium-dependent lipid-binding (CaLB domain) family protein
ath-miR414	miR414	AT5G64830	A_thaliana_only	0.676	Cat_1	702.06	UCAUCUUCAUCAUCAUCGUCA		programmed cell death 2 C-terminal domain-containing protein
ath-miR5658	miR5658	AT2G32310	A_thaliana_only	0.735	Cat_2	110.07	AUGAUGAUGAUGAUGAUGAAA		CCT motif family protein
ath-miR2933a	miR2933	AT4G32390	A_thaliana_only	0.765	Cat_1	20.78	GAAAUCGGAGAGGAAAUUCGCC		Nucleotide-sugar transporter family protein
ath-miR5027	miR5027	AT1G07610	A_thaliana_only	0.882	Cat_1	181.92	ACCGGUUGGAACUUGCCUUAA	MT1C	metallothionein 1C
ath-miR5652	miR5652	AT5G16640	A_thaliana_only	0.912	Cat_1	16.11	UUGAAUGUGAAUGAAUCGGGC		Pentatricopeptide repeat (PPR) superfamily protein
ath-miR5633	miR5633	AT2G35670	A_thaliana_only	1.147	Cat_1	117.11	UAUGAUCAUCAGAAAACAGUG	FIS2	VEFS-Box of polycomb protein
ath-miR5658	miR5658	AT5G56860	A_thaliana_only	1.382	Cat_1	100.43	AUGAUGAUGAUGAUGAUGAAA	GNC	GATA type zinc finger transcription factor family protein
ath-miR5658	miR5658	AT4G11600	A_thaliana_only	1.5	Cat_1	109.77	AUGAUGAUGAUGAUGAUGAAA	GPX6	glutathione peroxidase 6
ath-miR5658	miR5658	AT1G73710	A_thaliana_only	1.706	Cat_1	109.77	AUGAUGAUGAUGAUGAUGAAA		Pentatricopeptide repeat (PPR) superfamily protein
ath-miR414	miR414	AT1G60220	A_thaliana_only	1.853	Cat_1	702.06	UCAUCUUCAUCAUCAUCGUCA	ULP1D	UB-like protease 1D
ath-miR414	miR414	AT1G16150	A_thaliana_only	2.088	Cat_1	1404.12	UCAUCUUCAUCAUCAUCGUCA	WAKL4	wall associated kinase-like 4
ath-miR414	miR414	AT3G11810	A_thaliana_only	2.118	Cat_1	702.06	UCAUCUUCAUCAUCAUCGUCA		(1 of 2) PTHR33133:SF7 - F26K24.10 PROTEIN
ath-miR414	miR414	AT5G55300	A_thaliana_only	2.118	Cat_1	702.06	UCAUCUUCAUCAUCAUCGUCA	TOP1ALPHA	DNA topoisomerase I alpha
ath-miR5652	miR5652	AT1G62670	A_thaliana_only	2.529	Cat_1	38.52	UUGAAUGUGAAUGAAUCGGGC	RPF2	rna processing factor 2
ath-miR8177	miR8177	AT1G15710	A_thaliana_only	2.618	Cat_1	701.17	GUGUGAUGAUGUGUCAUUUAUA		prephenate dehydrogenase family protein
ath-miR414	miR414	AT5G40340	A_thaliana_only	2.765	Cat_1	1164.06	UCAUCUUCAUCAUCAUCGUCA		Tudor/PWWP/MBT superfamily protein
ath-miR5021	miR5021	AT3G23890	A_thaliana_only	3.559	Cat_1	776.04	UGAGAAGAAGAAGAAGAAAA	TOPII	topoisomerase II
ath-miR5021	miR5021	AT1G03190	A_thaliana_only	3.647	Cat_1	388.02	UGAGAAGAAGAAGAAGAAAA	UVH6	RAD3-like DNA-binding helicase protein
ath-miR5021	miR5021	AT2G40520	A_thaliana_only	3.676	Cat_1	388.02	UGAGAAGAAGAAGAAGAAAA		Nucleotidyltransferase family protein

ath-miR5021	miR5021	AT5G24670	A_thaliana_only	3.676	Cat_1	1164.06	UGAGAAGAAGAAGAAGAAAA		Cytidine/deoxycytidylate deaminase family protein
ath-miR414	miR414	AT5G55580	A_thaliana_only	3.941	Cat_1	702.06	UCAUCUUCAUCAUCAUCGUCA		Mitochondrial transcription termination factor family protein
ath-miR5029	miR5029	No HE Targets	A_thaliana_only	NA	NA	NA	AAUGAGAGAGAACACUGCAAA	NA	NA
ath-miR5595a	miR5595	No HE Targets	A_thaliana_only	NA	NA	NA	ACAUAUGAUCUGCAUCUUUGC	NA	NA
ath-miR5628	miR5628	No HE Targets	A_thaliana_only	NA	NA	NA	GAAAUAGCGAAGAUAUGAUUA	NA	NA
ath-miR5632-3p	miR5632	No HE Targets	A_thaliana_only	NA	NA	NA	UUGGAUUUAUAGUUGGAUAAG	NA	NA
ath-miR5634	miR5634	No HE Targets	A_thaliana_only	NA	NA	NA	AGGGACUUUGUGAAUUUAGGG	NA	NA
ath-miR5635a	miR5635	No HE Targets	A_thaliana_only	NA	NA	NA	UGUUAAGGAGUGUUAACGGUG	NA	NA
ath-miR5637	miR5637	No HE Targets	A_thaliana_only	NA	NA	NA	AAUGCGCAACUCUAUAUUUCC	NA	NA
ath-miR5638a	miR5638	No HE Targets	A_thaliana_only	NA	NA	NA	AUACCAAAACUCUCUCACUUU	NA	NA
ath-miR1886.2	miR1886	No HE Targets	A_thaliana_only	NA	NA	NA	UGAGAUGAAAUCUUUGAUUGG	NA	NA
ath-miR1888a	miR1888	No HE Targets	A_thaliana_only	NA	NA	NA	UAAGUUAAGAUUUGUGAAGAA	NA	NA
ath-miR2934-5p	miR2934	No HE Targets	A_thaliana_only	NA	NA	NA	UCUUUCUGCAAACGCCUUGGA	NA	NA
ath-miR2936	miR2936	No HE Targets	A_thaliana_only	NA	NA	NA	CUUGAGAGAGAGAACACAGACG	NA	NA
ath-miR2938	miR2938	No HE Targets	A_thaliana_only	NA	NA	NA	GAUCUUUUGAGAGGGUUCCAG	NA	NA
ath-miR2939	miR2939	No HE Targets	A_thaliana_only	NA	NA	NA	UAACGCACAACACUAAGCCAU	NA	NA
ath-miR3932a	miR3932	No HE Targets	A_thaliana_only	NA	NA	NA	AACUUUGUGAUGACAACGAAG	NA	NA
ath-miR3933	miR3933	No HE Targets	A_thaliana_only	NA	NA	NA	AGAAGCAAAAUGACGACUCGG	NA	NA
ath-miR401	miR401	No HE Targets	A_thaliana_only	NA	NA	NA	CGAAACUGGUGUCGACCGACA	NA	NA
ath-miR404	miR404	No HE Targets	A_thaliana_only	NA	NA	NA	AUUAACGCUGGCGGUUGCGGCAGC	NA	NA
ath-miR405a	miR405	No HE Targets	A_thaliana_only	NA	NA	NA	AUGAGUUGGGUCUAACCCAUAACU	NA	NA
ath-miR407	miR407	No HE Targets	A_thaliana_only	NA	NA	NA	UUUAAAUCAUAUACUUUUGGU	NA	NA
ath-miR413	miR413	No HE Targets	A_thaliana_only	NA	NA	NA	AUAGUUUCUCUUGUUCUGCAC	NA	NA
ath-miR417	miR417	No HE Targets	A_thaliana_only	NA	NA	NA	GAAGGUAGUGAAUUUGUUCGA	NA	NA
ath-miR418	miR418	No HE Targets	A_thaliana_only	NA	NA	NA	UAAUGUGAUGAUGAACUGACC	NA	NA
ath-miR419	miR419	No HE Targets	A_thaliana_only	NA	NA	NA	UUAUGAAUGCUGAGGAUGUUG	NA	NA
ath-miR420	miR420	No HE Targets	A_thaliana_only	NA	NA	NA	UAAACUAAUCACGGAAAUGCA	NA	NA
ath-miR426	miR426	No HE Targets	A_thaliana_only	NA	NA	NA	UUUUGGAAAUUUGUCCUUACG	NA	NA
ath-miR447a.2-3p	miR447	No HE Targets	A_thaliana_only	NA	NA	NA	UAUGGAAGAAAUUGUAGUAUU	NA	NA
ath-miR5012	miR5012	No HE Targets	A_thaliana_only	NA	NA	NA	UUUUACUGCUACUUGUGUUCC	NA	NA
ath-miR5013	miR5013	No HE Targets	A_thaliana_only	NA	NA	NA	UUUGUGACAUCUAGGUGCUUU	NA	NA
ath-miR5014a-3p	miR5014	No HE Targets	A_thaliana_only	NA	NA	NA	UUGUACAAAUUUAAGUGUACG	NA	NA
ath-miR5015	miR5015	No HE Targets	A_thaliana_only	NA	NA	NA	UUGGUGUUAUGUGUAGUCUUC	NA	NA
ath-miR5016	miR5016	No HE Targets	A_thaliana_only	NA	NA	NA	UUCUUGUGGAUUCCUUGGAAA	NA	NA
ath-miR5017-3p	miR5017	No HE Targets	A_thaliana_only	NA	NA	NA	UUAUACCAAAUUAAUAGCAAA	NA	NA
ath-miR5018	miR5018	No HE Targets	A_thaliana_only	NA	NA	NA	UUAAAGCUCCACCAUGAGUCCAAU	NA	NA
ath-miR5019	miR5019	No HE Targets	A_thaliana_only	NA	NA	NA	UGUUGGGAAAGAAAAACUCUU	NA	NA

ath-miR5020b	miR5020	No HE Targets	A_thaliana_only	NA	NA	NA	AUGGCAUGAAAGAAGGUGAGA	NA	NA
ath-miR5023	miR5023	No HE Targets	A_thaliana_only	NA	NA	NA	AUUGGUAGUGGAUAAGGGGGC	NA	NA
ath-miR5025	miR5025	No HE Targets	A_thaliana_only	NA	NA	NA	ACUGUAUAUAUGUAAGUGACA	NA	NA
ath-miR5026	miR5026	No HE Targets	A_thaliana_only	NA	NA	NA	ACUCAUAAGAUCGUGACACGU	NA	NA
ath-miR5028	miR5028	No HE Targets	A_thaliana_only	NA	NA	NA	AAUUGGGUUUAUGCUAGAGUU	NA	NA
ath-miR8121	miR8121	No HE Targets	A_thaliana_only	NA	NA	NA	AAAGUAUAAUGGUUUAGUGGUUUG	NA	NA
ath-miR8165	miR8165	No HE Targets	A_thaliana_only	NA	NA	NA	AAUGGAGGCAAGUGUGAAGGA	NA	NA
ath-miR8166	miR8166	No HE Targets	A_thaliana_only	NA	NA	NA	AGAGAGUGUAGAAAGUUUCUCA	NA	NA
ath-miR8168	miR8168	No HE Targets	A_thaliana_only	NA	NA	NA	AGGUGCUGAGUGUGCUAGUGC	NA	NA
ath-miR8169	miR8169	No HE Targets	A_thaliana_only	NA	NA	NA	AUAGACAGAGUCACUCACAGA	NA	NA
ath-miR8170-3p	miR8170	No HE Targets	A_thaliana_only	NA	NA	NA	UUGCUUAAAGAUUUUCUAUGU	NA	NA
ath-miR8172	miR8172	No HE Targets	A_thaliana_only	NA	NA	NA	AUGGAUCAUCUAGAUGGAGAU	NA	NA
ath-miR8173	miR8173	No HE Targets	A_thaliana_only	NA	NA	NA	AUGUGCUGAUUCGAGGUGGGA	NA	NA
ath-miR5639-3p	miR5639	No HE Targets	A_thaliana_only	NA	NA	NA	UUUAGCCUCAGACCACGGUGGACU	NA	NA
ath-miR5640	miR5640	No HE Targets	A_thaliana_only	NA	NA	NA	UGAGAGAAGGAAUUAGAUUCA	NA	NA
ath-miR5641	miR5641	No HE Targets	A_thaliana_only	NA	NA	NA	UGGAAGAAGAUGAUAGAAUUA	NA	NA
ath-miR5642a	miR5642	No HE Targets	A_thaliana_only	NA	NA	NA	UCUCGCGCUUGUACGGCUUU	NA	NA
ath-miR5643a	miR5643	No HE Targets	A_thaliana_only	NA	NA	NA	AGGCUUUUAAGAUCUGGUUGC	NA	NA
ath-miR5645a	miR5645	No HE Targets	A_thaliana_only	NA	NA	NA	AUUUGAGUCAUGUCGUUAAG	NA	NA
ath-miR5646	miR5646	No HE Targets	A_thaliana_only	NA	NA	NA	GUUCGAGGCACGUUGGGAGG	NA	NA
ath-miR5647	miR5647	No HE Targets	A_thaliana_only	NA	NA	NA	UCAAGUUUGAUGACGAUUCCA	NA	NA
ath-miR5648-3p	miR5648	No HE Targets	A_thaliana_only	NA	NA	NA	AUCUGAAGAAAAUAGCGGCAU	NA	NA
ath-miR5649a	miR5649	No HE Targets	A_thaliana_only	NA	NA	NA	AUUGAAUAUGUUGGUUACUAU	NA	NA
ath-miR5655	miR5655	No HE Targets	A_thaliana_only	NA	NA	NA	AAGUAGACACAUAAGAAGGAG	NA	NA
ath-miR5656	miR5656	No HE Targets	A_thaliana_only	NA	NA	NA	ACUGAAGUAGAGAUUGGGUUU	NA	NA
ath-miR5659	miR5659	No HE Targets	A_thaliana_only	NA	NA	NA	CGAUGAAGGUCUUUGGAACGGUA	NA	NA
ath-miR5660	miR5660	No HE Targets	A_thaliana_only	NA	NA	NA	CAGGUGGUUAGUGCAAUGGAA	NA	NA
ath-miR5662	miR5662	No HE Targets	A_thaliana_only	NA	NA	NA	AGAGGUGACCAUUGGAGAUG	NA	NA
ath-miR5663-3p	miR5663	No HE Targets	A_thaliana_only	NA	NA	NA	UGAGAAUGCAAAUCCUUAGCU	NA	NA
ath-miR5664	miR5664	No HE Targets	A_thaliana_only	NA	NA	NA	AUAGUCAAUUUUAUCGGUCUG	NA	NA
ath-miR5665	miR5665	No HE Targets	A_thaliana_only	NA	NA	NA	UUGGUGGACAAGAUCUGGGAU	NA	NA
ath-miR5666	miR5666	No HE Targets	A_thaliana_only	NA	NA	NA	AUGGGACAUCGAGCAUUUAAU	NA	NA
ath-miR5995b	miR5995	No HE Targets	A_thaliana_only	NA	NA	NA	ACAUAUGAUCUGCAUCUUUGC	NA	NA
ath-miR5996	miR5996	No HE Targets	A_thaliana_only	NA	NA	NA	UGACAUCCAGAUAGAAGCUUUG	NA	NA
ath-miR5997	miR5997	No HE Targets	A_thaliana_only	NA	NA	NA	UGAAACCAAGUAGCUAAAUAG	NA	NA
ath-miR5998a	miR5998	No HE Targets	A_thaliana_only	NA	NA	NA	ACAGUUUGUGUUUUGUUUUGU	NA	NA
ath-miR5999	miR5999	No HE Targets	A_thaliana_only	NA	NA	NA	UCUUCACUAUUAGACGGACAA	NA	NA
ath-miR771	miR771	No HE Targets	A_thaliana_only	NA	NA	NA	UGAGCCUCUGUGGUAGCCCUCA	NA	NA
ath-miR773b-5p	miR773	No HE Targets	A_thaliana_only	NA	NA	NA	GGCAAUAACUUGAGCAAACA	NA	NA

ath-miR776	miR776	No HE Targets	A_thaliana_only	NA	NA	NA	UCUAAGUCUUCUAUUGAUGUU	NA	NA
ath-miR777	miR777	No HE Targets	A_thaliana_only	NA	NA	NA	UACGCAUUGAGUUUCGUUGCUU	NA	NA
ath-miR778	miR778	No HE Targets	A_thaliana_only	NA	NA	NA	UGGCUUGGUUUAUGUACACCG	NA	NA
ath-miR779.1	miR779	No HE Targets	A_thaliana_only	NA	NA	NA	UUCUGCUAUGUUGCUGCUCAU	NA	NA
ath-miR780.2	miR780	No HE Targets	A_thaliana_only	NA	NA	NA	UUCUUCGUGAAUAUCUGGCAU	NA	NA
ath-miR782	miR782	No HE Targets	A_thaliana_only	NA	NA	NA	ACAAACACCUUGGAUGUUCUU	NA	NA
ath-miR8182	miR8182	No HE Targets	A_thaliana_only	NA	NA	NA	UUGUGUUGCGUUUCUGUUGAUU	NA	NA
ath-miR8184	miR8184	No HE Targets	A_thaliana_only	NA	NA	NA	UUUGGUCUGAUUACGAAUGUA	NA	NA
ath-miR830-3p	miR830	No HE Targets	A_thaliana_only	NA	NA	NA	UAACUAUUUUGAGAAGAAGUG	NA	NA
ath-miR832-5p	miR832	No HE Targets	A_thaliana_only	NA	NA	NA	UGCUGGGAUCGGGAAUCGAAA	NA	NA
ath-miR836	miR836	No HE Targets	A_thaliana_only	NA	NA	NA	UCCUGUGUUUCCUUUGAUGCGUGG	NA	NA
ath-miR849	miR849	No HE Targets	A_thaliana_only	NA	NA	NA	UAACUAAACAUUGGUGUAGUA	NA	NA
ath-miR854a	miR854	No HE Targets	A_thaliana_only	NA	NA	NA	GAUGAGGAUAGGGAGGAGGAG	NA	NA
ath-miR855	miR855	No HE Targets	A_thaliana_only	NA	NA	NA	AGCAAAAGCUAAGGAAAAGGAA	NA	NA
ath-miR8174	miR8174	No HE Targets	A_thaliana_only	NA	NA	NA	AUGUGUAUAGGGAAGCUAAUC	NA	NA
ath-miR8175	miR8175	No HE Targets	A_thaliana_only	NA	NA	NA	GAUCCCCGGCAACGGCGCCA	NA	NA
ath-miR8176	miR8176	No HE Targets	A_thaliana_only	NA	NA	NA	GGCCGGUGGUCGCGAGAGGGA	NA	NA
ath-miR8178	miR8178	No HE Targets	A_thaliana_only	NA	NA	NA	UAACAGAGUAAUUGUACAGUG	NA	NA
ath-miR8179	miR8179	No HE Targets	A_thaliana_only	NA	NA	NA	UGACUGCAUUAACUUGAUCGU	NA	NA
ath-miR863-5p	miR863	No HE Targets	A_thaliana_only	NA	NA	NA	UUAUGUCUUGUUGAUCUCAAU	NA	NA
ath-miR864-5p	miR864	No HE Targets	A_thaliana_only	NA	NA	NA	UCAGGUAUGAUUGACUUCAAA	NA	NA
ath-miR866-5p	miR866	No HE Targets	A_thaliana_only	NA	NA	NA	UCAAGGAACGGAUUUUGUUAA	NA	NA
ath-miR867	miR867	No HE Targets	A_thaliana_only	NA	NA	NA	UUGAACAUGGUUUAUUAGGAA	NA	NA
ath-miR870-3p	miR870	No HE Targets	A_thaliana_only	NA	NA	NA	UAAUUUGGUGUUUCUUCGAUC	NA	NA

Table S7. IsomiRs of the conserved miRNAs used for analysis across species

Arabidopsis thaliana			**Amborella trichopoda**			**Brachypodium distachyon**	
miR156a	ugacagaagagagugagcac		miR156a	uugacagaagauagagagcac		miR156b	ugacagaagagagugagcac
miR159a	uuuggauugaagggagcucua		miR159	uuuggauugaagggagcucua		miR159b	uuuggauugaagggagcucu
miR160a	ugccuggcucccuguaugcca		miR160	ugccuggcucccuguaugcca		miR160a	ugccuggcucccuguaugcca
miR162a	ucgauaaaccucugcauccag		miR166a	ucggaccaggcuucauucccc		miR164a	uggagaagcagggcacgugca
miR164a	uggagaagcagggcacgugca		miR167	ugaagcugccagcaugaucug		miR166a	ucggaccaggcuucauucccc
miR166a	ucggaccaggcuucauucccc		miR169a	uagccaaggaugacuugccu		miR167a	ugaagcugccagcaugaucua
miR167a	ugaagcugccagcaugaucua		miR172a	ggaaucuugaugaugcugca		miR168	ucgcuuggugcagaucgggac
miR168a	ucgcuuggugcaggucgggaa		miR319a	uuggacugaagggagcuccc		miR169a	cagccaaggaugacuugccga
miR169a	cagccaaggaugacuugccga		miR393	uccaaagggaucgcauugaucc		miR171b	ugauugagccgcgccaauauc
miR171a	ugauugagccgcgccaauauc		miR394	uuggcauucuguccaccucc		miR172a	agaaucuugaugaugcugcau
miR172a	agaaucuugaugaugcugcau		miR395	cugaaguguuugggggaacuc		miR319b	uuggacugaagggugcucccu
miR319a	uuggacugaagggagcucccu		miR396b	uuccacagcuuucuugaacau		miR393a	uccaaagggaucgcauugauc
miR393a	uccaaagggaucgcauugaucc		miR398	uguguucccaggucgccccug		miR394	uuggcauucuguccaccucc
miR394a	uuggcauucuguccaccucc		TasiARFs	uucuugaccuuguaagaccuu		miR395	ugaaguguuugggggaacuc
miR395a	cugaaguguuugggggaacuc					miR396a	uccacaggcuuucuugaacug
miR396a	uuccacagcuuucuugaacug					miR398a	uguguucucaggucgccccug
miR397a	ucauugagugcagcguugaug					miR408	cugcacugccucuucccuggc
miR398b	uguguucucaggucaccccug					TasiARFs	uucuugaccuuguaagaccuu
miR403	uuagauucacgcacaaacucg						
miR408	augcacugccucuucccuggc						
TasiARFs	uucuugaccuuguaagaccuu						
Citrus sinensis			**Glycine max**			**Hordeum vulgare**	
miR156a	ugacagaagagagugagcac		miR156a	ugacagaagagagugagcac		miR156a	ugacagaagagagugagcaca
miR159a	uuuggauugaagggagcucua		miR159a	uuuggauugaagggagcucua		miR159a	uuuggauugaagggagcucug
miR162	ucgauaaaccucugcauccag		miR160a	gccuggcucccuguaugccau			
miR164a	uggagaagcagggcacgugca		miR162b	ucgauaaaccucugcauccag		**Malus domestica**	
miR166b	ucucggaccaggcuucauucc		miR164a	uggagaagcagggcacgugca		miR156a	ugacagaagagagugagcac
miR167a	gaagcugccagcaugaucug		miR166a	ucggaccaggcuucauucccc		miR159a	cuuggauugaagggagcucc
miR168	ucgcuuggugcaggucgggaa		miR167c	ugaagcugccagcaugaucug		miR162a	ucgauaaaccucugcauccag
miR171a	uugagccgugccaauaucac		miR168a	ucgcuuggugcaggucgggaa		miR164a	uggagaagcagggcacaugcc
miR172b	agaaucuugaugaugcugcau		miR169a	cagccaaggaugacuugccgg		miR166a	ucggaccaggcuucauucccc
miR395	cugaaguguuugggggaacuc		miR171c	uugagccgugccaauaucaca		miR168a	ucgcuuggugcaggucgggaa
miR396a	uuccacagcuuucuugaacug		miR172a	agaaucuugaugaugcugcau		miR172a	agaaucuugaugaugcugca
miR397	ucauugagugcagcguugaug		miR319a	uuggacugaagggagcuccc		miR319a	uuggacugaagggagcucccu
miR398b	uguguucucaggucgccccug		miR393h	uuccaaagggaucgcauugauc		miR393a	uccaaagggaucgcauugaucu
miR408	augcacugccucuucccuggc		miR394a	uuggcauucuguccaccucc		miR395a	cugaaguguuugggggaacuc

TasiARFs	uucuugaccuuguaagaccuu		miR395a	cugaaguguuugggggaacuc		miR398a	uguguucucaggucgccccug
			miR396a	uuccacagcuuucuugaacug		miR403a	uuagauucacgcacaaacucg
			miR397a	ucauugagugcagcguugaug		miR408a	augcacugccucuucccuggc
			miR403a	uuagauucacgcacaaacuug		TasiARFs	uucuugaccuuguaagaccuu
			miR408a	augcacugccucuucccuggc			
			TasiARFs	uucuugaccuuguaagaccuu			
Medicago truncatula			**Oryza sativa**			**Prunus persica**	
miR156b	ugacagaagagagugagcac		miR156a	ugacagaagagagugagcac		miR156a	ugacagaagaaagagagcac
miR159a	uuuggauugaagggagcucua		miR159a	uuuggauugaagggagcucug		miR159	uuuggauugaagggagcucua
miR160a	ugccuggcucccuguaugcca		miR160a	ugccuggcucccuguaugcca		miR160a	ugccuggcucccuguaugcca
miR162	ucgauaaaccucugcauccag		miR162a	ucgauaaaccucugcauccag		miR164a	uggagaagcagggcacgugca
miR164a	uggagaagcagggcacgugca		miR164a	uggagaagcagggcacgugca		miR166a	ucggaccaggcuucauucccc
miR166a	ucggaccaggcuucauucccc		miR166a	ucggaccaggcuucauucccc		miR167a	ugaagcugccagcaugaucua
miR167a	ugaagcugccagcaugaucua		miR167a	ugaagcugccagcaugaucua		miR168	ucgcuuggugcaggucgggaa
miR169a	cagccaaggaugacuugccga		miR168a	ucgcuuggugcagaucgggac		miR169a	cagccaaggaugacuugccgg
miR171a	ugauugagucgugccaauauc		miR169a	cagccaaggaugacuugccga		miR171a	ugauugagccgugccaauauc
miR172a	agaauccugaugaugcugcag		miR171b	ugauugagccgugccaauauc		miR172a	agaaucuugaugaugcugcau
miR319a	uuggacugaagggagcuccc		miR172a	agaaucuugaugaugcugcau		miR319a	uuggacugaagggagcuccc
miR393a	uccaaagggaucgcauugauc		miR319b	uuggacugaagggugcuccc		miR393a	uccaaagggaucgcauugauc
miR396a	uuccacagcuuucuugaacuu		miR393a	uccaaagggaucgcauugauc		miR394a	uuggcauucuguccaccucc
miR397	ucauugagugcagcguugaug		miR394	uuggcauucuguccaccucc		miR395a	cugaaguguuugggggaccc
miR398b	uguguucucaggucgccccug		miR396e	uccacaggcuuucuugaacug		miR396a	uuccacagcuuucuugaacgu
miR408	augcacugccucuucccuggc		miR397a	ucauugagugcagcguugaug		miR397a	ucauugagugcagcguugaug
TasiARFs	uucuugaccuuguaagaccuu		miR398b	uguguucucaggucaccccug		miR398a	uguguucucaggucgccccug
			miR408	cugcacugccucuucccuggc		miR403	uuagauucacgcacaaacucg
			TasiARFs	uucuugaccuuguaagaccuu		TasiARFs	uucuugaccuuguaagaccuu
Solanum lycopersicum			**Selanginella moellendorffii**			**Vitis vinifera**	
miR156a	uugacagaagauagagagcac		miR156a	cgacagaagagagugagcac		miR156b	ugacagaagagagugagcac
miR159	uuuggauugaagggagcucua		miR160	ugccuggcucccuguaugcca		miR159a	cuuggagugaagggagcucuc
miR160a	ugccuggcucccuguaugcca		miR166a	ucggaccaggcuucauucccc		miR160	ugccuggcucccuguaugcca
miR164a	uggagaagcagggcacgugca		miR319	ugcugccgacucaugcaucc		miR162	ucgauaaaccucugcauccag
miR166a	ucggaccaggcuucauucccc		miR396a	uuccacggcuuucuugaacc		miR164a	uggagaagcagggcacgugca
miR167a	ugaagcugccagcaugaucua					miR166a	ucggaccaggcuucauucccc
miR168a	ucgcuuggugcaggucgggac		**Triticum aestivum**			miR169a	cagccaaggaugacuugccgg
miR169a	cagccaaggaugacuugccgg		miR156	ugacagaagagagugagcaca		miR171a	ugauugagccgugccaauauc
miR171a	ugauugagccgugccaauauc		miR159a	uuuggauugaagggagcucug		miR172a	ugaaucuugaugaugcuacau
miR172a	agaaucuugaugaugcugcau		miR160	ugccuggcucccuguaugcca		miR319b	uuggacugaagggagcucccu
miR319b	uuggacugaagggagcucccu		miR164a	uggagaagcagggcacgugca		miR395a	cugaaguguuugggggaacuc
miR394	uuggcauucuguccaccucc		miR167a	ugaagcugccagcaugaucua		miR396a	uuccacagcuuucuugaacua

miR395a	cugaaguguuugggggaacucc		miR395a	gugaaguguuugggggaacuc		miR398b	uguguucucaggucaccccug
miR396a	uuccacagcuuucuugaacug		miR398	uguguucucaggucgcccccg		miR403a	uuagauucacgcacaaacucg
miR397	auugagugcagcguugauga					miR408	augcacugccucuucccuggc
miR398a	uauguucucaggucgccccug					TasiARFs	uucuugaccuuguaagaccuu
miR403	cuagauucacgcacaagcucg						
TasiARFs	uucuugaccuuguaagaccuu						
Zea mays							
miR156a	ugacagaagagagugagcac						
miR159a	uuuggauugaagggagcucug						
miR160a	ugccuggcucccuguaugcca						
miR162	ucgauaaaccucugcaucca						
miR164a	uggagaagcagggcacgugca						
miR166a	ucggaccaggcuucauucccc						
miR167a	ugaagcugccagcaugaucua						
miR168a	ucgcuuggugcagaucgggac						
miR169a	cagccaaggaugacuugccga						
miR171d	ugauugagccgugccaauauc						
miR172a	agaaucuugaugaugcugca						
miR319a	uuggacugaagggugcuccc						
miR393a	uccaaagggaucgcauugaucu						
miR394a	uuggcauucuguccaccucc						
miR396a	uuccacagcuuucuugaacug						
miR397a	ucauugagcgcagcguugaug						
miR398a	uguguucucaggucgcccccg						
miR408a	cugcacugccucuucccuggc						
TasiARFs	uucuugaccuuguaagaccuu						

Table S8. Transcriptome libraries used for psRNATarget and WPMIAS

Species	**psRNATarget Library***	**WPMIAS Library****
Dicotyledons		
Arabidopsis thaliana	transcript, JGI genomic project, Phytozome 11, 167_TAIR10	Transcript, JGI genomic project, Phytozome 11, 167 TAIR10(from psRNAtarget)
Citrus sinensis	transcript, JGI genomic project, Phytozome 13, 154_v1.1	transcript, JGI genomic project, Phytozome 11, 154_v1.1(from psRNAtarget)
Glycine max	transcript, JGI genomic project, Phytozome 11, 275	transcript, JGI genomic project, Phytozome 12, Wm82.a2.v1
Malus domestica	JGI genomic project, Phytozome 11, 196 v1.0	Transcript, Phytozome v12, released on 2014/1/8
Medicago truncatula	transcript, JGI genomic project, Phytozome 11, 285 Mt4.0 v1	transcript, Mt4.0v1 spliced transcripts, IMGAG, Mt4.0V1
Prunus persica	transcript, JGI genomic project, Phytozome 13, 298, v2.1	Transcript, Phytozome v12, released on 2014/1/8
Solanum lycopersicum	JGI genomic project, Phytozome 12, 390_ITAG2.4	JGI genomic project, Phytozome 11, 390_ITAG2.4
Vitis vinifera	transcript, JGI genomic project, Phytozome 12, 145_Genoscope.12X	transcript, JGI genomic project, Phytozome 11, 145 Genoscope 12X(from psRNAtarget)
Monocotyledons		
Zea mays	transcript, JGI genomic project, Phytozome 11, Ensembl-18_2010_01	transcript, JGI genomic project, Phytozome 12, Ensembl-18_2010_01
Brachypodium distachyon	transcript, JGI genomic project, Phytozome 11, 314_v3.1	Transcript, Phytozome, v12, released on 2015/8/25
Hordeum vulgare	cDNA, EnsemblePlants library, Hv_IBSC_PGSB_v2, 2018	cDNA, EnsemblePlants library, Hv_IBSC_PGSB_v2, 2018
Oryza sativa	transcript, JGI genomic project, Phytozome 13, 323_v7.0	Transcript, Phytozome v12, released on 2015/11/27
Triticum aestivum	transcript, cDNA library, TGACv1	cDNA, EnsemblePlants library, TGACv1.release39, 2018
Amborella trichopoda	transcript, JGI genomic project, Phytozome 13, 291_v1.0	Transcript, JGI genomic project, Phytozome 11, 291_v1.0(from psRNAtarget)
Selaginella moellendorffii	transcript, JGI genomic project, Phytozome 11, 91_v1.0	transcript, JGI genomic project, Phytozome 12, 91_v1.0
*Transcriptome libraries used for psRNATarget		
**Transcriptome libraries used for WPMIAS		

Table S9. Transcriptomes used for identifying conserved sequences in target transcripts

Transcriptomes downloaded from the Genome portal of the Department of Energy Joint Genome Institute (Grigoriev et al., 2012; Nordberg et al., 2014) (https://phytozome.jgi.doe.gov/pz/portal.html)		
	Transcriptome file name	**Reference**
Species		
Dicotyledons		
Arabidopsis thaliana	Athaliana_167_TAIR10.transcript_primaryTranscriptOnly.fa	Lamesch, P., Berardini, T. Z., Li, D., Swarbreck, D., Wilks, C., Sasidharan, R., Muller, R., Dreher, K., Alexander, D. L., Garcia-Hernandez, M., Karthikeyan, A. S., Lee, C. H., Nelson, W. D., Ploetz, L., Singh, S., Wensel, A., & Huala, E. (2012). The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. Nucleic acids research, 40(Database issue), D1202–D1210. https://doi.org/10.1093/nar/gkr1090
Citrus sinensis	Csinensis_154_transcript	Wu, G. A., Prochnik, S., Jenkins, J., Salse, J., Hellsten, U., Murat, F., Perrier, X., Ruiz, M., Scalabrin, S., Terol, J., Takita, M. A., Labadie, K., Poulain, J., Couloux, A., Jabbari, K., Cattonaro, F., Del Fabbro, C., Pinosio, S., Zuccolo, A., Chapman, J., … Rokhsar, D. (2014). Sequencing of diverse mandarin, pummelo and orange genomes reveals complex history of admixture during citrus domestication. Nature biotechnology, 32(7), 656–662. https://doi.org/10.1038/nbt.2906
Glycine max	Gmax_275_Wm82.a2.v1.transcript.fa	Schmutz, J., Cannon, S. B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., Hyten, D. L., Song, Q., Thelen, J. J., Cheng, J., Xu, D., Hellsten, U., May, G. D., Yu, Y., Sakurai, T., Umezawa, T., Bhattacharyya, M. K., Sandhu, D., Valliyodan, B., Lindquist, E., … Jackson, S. A. (2010). Genome sequence of the palaeopolyploid soybean. Nature, 463(7278), 178–183. https://doi.org/10.1038/nature08670
Malus domestica	Mdomestica_196_v1.0.transcript	Velasco, R., Zharkikh, A., Affourtit, J., Dhingra, A., Cestaro, A., Kalyanaraman, A., Fontana, P., Bhatnagar, S. K., Troggio, M., Pruss, D., Salvi, S., Pindo, M., Baldi, P., Castelletti, S., Cavaiuolo, M., Coppola, G., Costa, F., Cova, V., Dal Ri, A., Goremykin, V., … Viola, R. (2010). The genome of the domesticated apple (Malus × domestica Borkh.). Nature genetics, 42(10), 833–839. https://doi.org/10.1038/ng.654
Medicago truncatula	Mtruncatula_285_Mt4.0v1.transcript	Tang, H., Krishnakumar, V., Bidwell, S., Rosen, B., Chan, A., Zhou, S., Gentzbittel, L., Childs, K. L., Yandell, M., Gundlach, H., Mayer, K. F., Schwartz, D. C., & Town, C. D. (2014). An improved genome release (version Mt4.0) for the model legume Medicago truncatula. BMC genomics, 15, 312. https://doi.org/10.1186/1471-2164-15-312
Prunus persica	Ppersica_298_v2.1.transcript	International Peach Genome Initiative, Verde, I., Abbott, A. G., Scalabrin, S., Jung, S., Shu, S., Marroni, F., Zhebentyayeva, T., Dettori, M. T., Grimwood, J., Cattonaro, F., Zuccolo, A., Rossini, L., Jenkins, J., Vendramin, E., Meisel, L. A., Decroocq, V., Sosinski, B., Prochnik, S., Mitros, T., … Rokhsar, D. S. (2013). The high-quality draft genome of peach (Prunus persica) identifies unique patterns of genetic diversity, domestication and genome evolution. Nature genetics, 45(5), 487–494. https://doi.org/10.1038/ng.2586
Solanum lycopersicum	Slycopersicum_390_ITAG2.4.transcript	Tomato Genome Consortium (2012). The tomato genome sequence provides insights into fleshy fruit evolution. Nature, 485(7400), 635–641. https://doi.org/10.1038/nature11119
Vitis vinifera	Vvinifera_145_Genoscope.12X.transcript	Jaillon, O., Aury, J. M., Noel, B., Policriti, A., Clepet, C., Casagrande, A., Choisne, N., Aubourg, S., Vitulo, N., Jubin, C., Vezzi, A., Legeai, F., Hugueney, P., Dasilva, C., Horner, D., Mica, E., Jublot, D., Poulain, J., Bruyère, C., Billault, A., … French-Italian Public Consortium for Grapevine Genome Characterization (2007). The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature, 449(7161), 463–467. https://doi.org/10.1038/nature06148
Monocotyledons		
Zea mays	Zmays_284_Ensembl-18_2010-01-MaizeSequence.transcript	Schnable, P. S., Ware, D., Fulton, R. S., Stein, J. C., Wei, F., Pasternak, S., Liang, C., Zhang, J., Fulton, L., Graves, T. A., Minx, P., Reily, A. D., Courtney, L., Kruchowski, S. S., Tomlinson, C., Strong, C., Delehaunty, K., Fronick, C., Courtney, B., Rock, S. M., … Wilson, R. K. (2009). The B73 maize genome: complexity, diversity, and dynamics. Science (New York, N.Y.), 326(5956), 1112–1115. https://doi.org/10.1126/science.1178534
Brachypodium distachyon	Bdistachyon_314_v3.1.transcript	International Brachypodium Initiative (2010). Genome sequencing and analysis of the model grass Brachypodium distachyon. Nature, 463(7282), 763–768. https://doi.org/10.1038/nature08747

Hordeum vulgare	Hvulgare_462_r1.transcript	Beier, S., Himmelbach, A., Colmsee, C., Zhang, X. Q., Barrero, R. A., Zhang, Q., Li, L., Bayer, M., Bolser, D., Taudien, S., Groth, M., Felder, M., Hastie, A., Šimková, H., Staňková, H., Vrána, J., Chan, S., Muñoz-Amatriaín, M., Ounit, R., Wanamaker, S., … Mascher, M. (2017). Construction of a map-based reference genome sequence for barley, Hordeum vulgare L. Scientific data, 4, 170044. https://doi.org/10.1038/sdata.2017.44
Oryza sativa	Osativa_323_v7.0.transcript	Ouyang, S., Zhu, W., Hamilton, J., Lin, H., Campbell, M., Childs, K., Thibaud-Nissen, F., Malek, R. L., Lee, Y., Zheng, L., Orvis, J., Haas, B., Wortman, J., & Buell, C. R. (2007). The TIGR Rice Genome Annotation Resource: improvements and new features. Nucleic acids research, 35(Database issue), D883–D887. https://doi.org/10.1093/nar/gkl976
Amborella trichopoda	Atrichopoda_291_v1.0.transcript	Amborella Genome Project (2013). The Amborella genome and the evolution of flowering plants. Science (New York, N.Y.), 342(6165), 1241089. https://doi.org/10.1126/science.1241089
Selaginella moellendorffii	Smoellendorffii_91_v1.0.transcript	Banks, J. A., Nishiyama, T., Hasebe, M., Bowman, J. L., Gribskov, M., dePamphilis, C., Albert, V. A., Aono, N., Aoyama, T., Ambrose, B. A., Ashton, N. W., Axtell, M. J., Barker, E., Barker, M. S., Bennetzen, J. L., Bonawitz, N. D., Chapple, C., Cheng, C., Correa, L. G., Dacre, M., … Grigoriev, I. V. (2011). The Selaginella genome identifies genetic changes associated with the evolution of vascular plants. Science (New York, N.Y.), 332(6032), 960–963. https://doi.org/10.1126/science.1203810

Table S10. Primers

Name	**Template**	**Sequence 5'-3'**	**Purpose**
ARF10_5UTR_FP	Arabidopsis genomic DNA	GGGGACAAGTTTGTACAAAAAAGCAGGCTATGGAGCAAGAGAAAAGC	*ARF10-WT* gateway construct
ARF10_3UTR_RP	Arabidopsis genomic DNA	GGGGACCACTTTGTACAAGAAAGCTGGGTGTTACACACAAAAGACCAGC	*ARF10-WT* gateway construct
ARF10-FM_SDM_FP	ARF10-WT (pDONR/Zeo) entry vector	AATACAGGGAGCCAGGCAGGCGCAGCAACTCTTCGG	Site directed mutagenesis to create *ARF10-FM* entry clone
ARF10-FM_SDM_RP	ARF10-WT (pDONR/Zeo) entry vector	TGCCTGGCTCCCTGTATTCCTGCGGGTGCATTATTGTTG	Site directed mutagenesis to create *ARF10-FM* entry clone
rmARF10_FP_SDM	ARF10-WT (pDONR/Zeo) entry vector	ATTCAAGGGGCCCGACAAGCTCAACAACTCTTCGGATCACCATC	Site directed mutagenesis to create the rm*ARF10* entry clone
rmARF10-RP_SDM	ARF10-WT (pDONR/Zeo) entry vector	TGTCGGGCCCCTTGAATCCCTGCAGGAGCATTATTGTTGTCG	Site directed mutagenesis to create the rm*ARF10* entry clone