Functions and target selection of *Arabidopsis* microRNAs

der Fakultät für Biologie

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines Doktors

der Naturwissenschaften

von

Rebecca Schwab

aus Besigheim

vorgelegte

Dissertation

2006

Acknowledgements

Publications

During the course of this work, the following articles have been published or submitted for publication:

Palatnik, J.F., Allen, E., Wu, X., Schommer, C., Schwab, R., Carrington, J.C., and Weigel, D. (2003). Control of leaf morphogenesis by microRNAs. Nature **425,** 257-263.

Schwab, R., Palatnik, J.F., Riester, M., Schommer, C., Schmid, M., and Weigel, D. (2005). Specific effects of microRNAs on the plant transcriptome. Dev. Cell **8,** 517-527.

Schwab, R., Ossowski, S., Riester, M., Warthmann, N., and Weigel, D. (2006). Artificial microRNAs silence target genes with high specificity. Submitted for publication to Plant Cell.

Palatnik, J.F., Wollman, H., Schommer, C., Schwab, R., Boisbouvier, J., Allen, E., Dezulian, T., Huson, D., Carrington, J.C., and Weigel, D. (2006). Specificity determinants of two sequence-related plant miRNAs. Submitted for publication to PLoS Biology.

Table of contents

8

Summary

Spatial and temporal control of gene activity is a prerequisite of multicellular development, as the local accumulation of gene products is required to specify different cell fates. The regulation of gene expression involves a large number of different effectors, and also single-stranded RNAs of ~19-25 nucleotides in length. MicroRNAs (miRNAs) constitute a large group of endogenous small RNAs, and they negatively regulate their target genes by base-pairing to complementary nucleic acids. In animals, miRNAs typically trigger translational arrest of their targets, to which they pair with only limited complementarity. Plant miRNAs, on contrary, trigger cleavage of target transcripts, with which they share high sequence complementarity. Predictions of new plant miRNA targets have therefore focused on genes with only a small number of mismatches to miRNAs.

Studying the biological function of three plant miRNAs, I have over- and misexpressed their precursors in *Arabidopsis thaliana*. The different abnormalities, which could be observed in transgenic plants, strongly suggest a role of these miRNAs as regulators of plant development. Overexpression of miR156 extended the vegetative phase of *Arabidopsis* seedlings, and also increased the speed and number of organ initiation events. Conversely, miR172 was shown to decrease the time of vegetative development when overexpressed, and also to control floral organ identity. MiRNA164 was shown to control separation of above ground organs throughout the life cycle of *Arabidopsis* plants. The observed defects can be partially correlated with known loss-of-function mutants in predicted target genes, however they also postulate additional miRNA functions. These might be mediated by other predicted target genes, which had not been functionally characterized before.

I have studied direct effects of miRNA expression on target RNA accumulation and confirmed that many predicted target genes were strongly responsive to miRNA overaccumulation. This finding is consistent with phenotypes of miRNA overexpressers resembling mutants in target genes, and also with cleavage of mRNAs as the mode of plant miRNA function. In order to test for the presence of additional targets with similar or lower sequence complementarity, I monitored genome-wide expression changes caused by overexpression of five different miRNAs. In brief, these analyses suggest that plant miRNAs directly regulate only a very small number of target genes, to which they pair with high sequence complementarity.

Comparing authentic miRNA targets to other, non-responsive genes with similar mismatch numbers, I established positional determinants of plant miRNA target selection. These, unlike previous computational efforts, almost unambiguously discriminate targets from non-targets not only in the small set of miRNAs analyzed, but also when compared to other plant miRNAs.

Feedback regulation can add another level of complexity to miRNA mediated effects, as I have discovered for the case of miR172. It has been previously suggested that miR172 overexpression results in changes of target protein, rather than transcript abundance, and translational inhibition was proposed as a mechanistic basis. However, since cleavage of target transcripts was efficiently increased in miR172 overexpressing tissue, additional regulatory events have to be present. To uncouple miRNA mediated from other effects on target genes, I overexpressed a miRNA resistant version of the target gene *AP2*, and confirmed the presence of a negative feedback of AP2 on its own expression. This suggests that miR172 functions by simultaneously triggering cleavage and translational inhibition of its target genes.

Since the action spectrum of plant miRNAs is very narrow, it contrasts with the broad selectivity of animal miRNAs. This difference might reflect either only intrinsic properties of the plant miRNA machinery, or selection against miRNAs with broader specificity has reduced the number of plant miRNA targets. In order to distinguish between the two possibilities, I have generated artificial miRNAs (amiRNAs) targeting endogenous genes and found that their specificity was as high as that of natural plant miRNAs. This finding supports the idea that extensive base-pairing with target genes is required for plant miRNA function. Since amiRNAs were efficiently produced and could be designed to specifically silence single, or groups of endogenous genes, they can easily be used as a tool for directed gene silencing in plants. In addition to conventional silencing of single genes, which is already possible by RNA interference, amiRNAs can specifically downregulate expression of multiple related genes. As they function with high specificity, they can potentially also function in strand or allele specific gene silencing, which is not possible by other means. Furthermore, introduction of amiRNA-insensitive variants of targets can be generates to compensate for defects in amiRNA expressing plants. A web-based tool has been established to automatically design amiRNAs for genes of interest and is available to the scientific community for further studies.

Introduction

Gene regulation and development

Growth and development of multi-cellular organisms is characterized by the specification and differentiation of diverse cell types and organs. Specification processes are based on a range of proteins and other molecules, and especially on their presence in specific places, at specific times, and in adequate concentrations. Therefore, their abundance needs to be tightly coordinated in both space and time.

Proteins are synthesized from nucleic acid templates, and the formation of other molecules is normally catalyzed by diverse proteins, making protein production and stability an effective step to regulate differentiation processes. Synthesis of proteins from their corresponding genes in the genome, is summarized as gene expression and comprises two spatially separated processes, which are regulated at several levels (Alberts et al. 2002). Genes are described as stretches of the nuclear DNA located on the chromosomes, and they are transcribed into single-stranded mobile RNA transcripts (mRNAs). These are exported into the cytoplasm, where they specify the sequence of polypeptide chains, the proteins. Regulatory steps act on the rate of transcription, transcript stability, efficiency of translation as well as protein stability.

Chromatin state greatly influences the rate of transcription, as it specifies accessibility of regulatory elements, normally preceding the protein coding elements of a gene. Heterochromatic regions, which are often found in centromeric regions, contain tightly packed DNA and histones and are normally transcriptionally silent, while transcriptionally active genes are mostly found in euchromatic regions, where packaging is less dense. Here, so-called transcription factors of either repressive or activating character can bind to regulatory elements and modify transcription states of the respective genes. This regulation is depending on the presence and/or activation state of these transcription factors and therefore state- and/or tissue-dependent.

Primary transcripts are processed in the nucleus and subsequently exported into the cytoplasm, where protein synthesis takes place. Stability of transcripts is largely determined by intramolecular structures and sequences, such as 3' poly-adenosine tails or 5' guanosine-caps, which are present on all transcripts generated by polymerase II. But also additional features, such as stable loop structures, are bound by specialized proteins, which stabilize or destabilize the corresponding transcripts. Sequence motifs within transcripts can be recognized by protein

complexes with the help of pairing nucleic acids and trigger their destruction. Alternatively, these protein complexes can hinder efficient translation of these transcripts. Proteins themselves are of different half-life, depending on sequence motifs, which can be recognized by other proteins, and trigger their destruction.

In addition to these various ways to regulate gene expression, individual steps are often cross-related and feedback regulation adds another level of complexity.

Small RNAs and gene regulation

Many levels of gene regulation can be influenced by different classes of small RNAs, which are single stranded RNA molecules of ~19-25 nucleotides in length (Bartel 2004). They are produced from double stranded precursors in both plants and animals and bind to target nucleic acids by complementary base-pairing. Based on differences in their biogenesis and action, small RNAs with regulatory functions in gene expression have been grouped into different classes, the most prominent being short interfering (si) RNAs and micro (mi) RNA.

The first **miRNA**, *lin-4* of 21 nucleotide in length, has been described in *C.elegans*, and was shown to inhibit translation of the heterochronic gene *lin-14*, with which it shares short elements of partial sequence complementarity in its 3'UTR (Lee et al. 1993; Wightman et al. 1993). *lin-14* mutants skip larval stages, while opposite phenotypes have been observed in *lin-4* mutants (reiteration of larval stages; (Horvitz and Sulston 1980), consistent with the notion, that the small RNA represses translation of *lin-14*, probably by binding to its 3' UTR. More than seven years later, miRNAs have also been identified from many other model organisms such as *Drosophila melanogaster* (Aravin et al. 2003), mouse and human (Lagos-Quintana et al. 2003; Lagos-Quintana et al. 2002) and also from various plant species (Billoud et al. 2005; Jones-Rhoades and Bartel 2004; Reinhart et al. 2002; Wang et al. 2004a; Zhang et al. 2005), and many of them regulate different aspects of development. While inhibition of translation is the main mode of miRNA action in animals, plant miRNAs typically trigger cleavage of target RNAs they bind to.

Short interfering RNAs of 24-25nt have been discovered in tobacco plants, which had been modified to express viral or transgene RNAs (Hamilton and Baulcombe 1999). Small RNAs of both sense and antisense orientation to the long exogenous RNA accumulated as a consequence of high levels of this RNAs. Coupled to down-regulation of the longer exogenous RNAs suggests that small

RNAs served as a silencing trigger (Hamilton and Baulcombe 1999). Independently, it has been found in *Caenothabditis elegans*, that the engineered presence of double stranded RNA leads to most efficient silencing of homologous transcripts, more efficiently than sense or antisense RNAs alone (Fire et al. 1998). As target transcript regulation occurs at the post-transcriptional level it has been generally termed post-transcriptional gene silencing (PTGS) (Montgomery et al. 1998). The involvement of small RNAs in this process was once more demonstrated by Zamore and colleagues (Zamore et al. 2000), showing that double stranded RNA silencing triggers are processed into ~21nt regularly spaced RNAs.

Gene silencing by siRNAs has been proposed to be sequence specific, as e.g. small RNAs corresponding to a green fluorescent protein (GFP) transgene can efficiently reduce levels of the GFP transcript, but do not influence levels of a second transgene (Voinnet et al. 1998). Additionally, it was shown that silencing of this GFP transgene does not only occur locally in infected plant cells, but can spread throughout the organism and cause systemic silencing of the respective transcript (Voinnet and Baulcombe 1997). Systemic silencing has been proposed to confer acquired immunity to viral infection in non-infected organs prior to the actual infection. Supporting this hypothesis, small RNAs have been found in the phloem sap of different plants and can also move across plasmodesmata, which symplastically connect most plant cells (Yoo et al. 2004).

A third mechanism of small RNA (mostly siRNA) action, has been described as transcriptional gene silencing (TGS), as it interferes with gene expression at the level of transcription. Small RNAs trigger methylation of homologous DNA, which is often coupled to methylation of histones in the corresponding chromatin, thus causing silencing of the respective locus. These methylation triggers are wide spread in silencing of repeats and transposons in different organisms (Gendrel and Colot 2005).

Small RNA biogenesis.

Small RNAs are synthesized from longer double stranded (ds) RNAs, which are processed by specialized ribonucleases into small double-stranded intermediary products of ~19-25 nucleotides in length (Zamore et al. 2000). These products typically contain 2 nucleotide overhangs at their 3' ends and free 5' phosphates (Elbashir et al. 2001a). The strands with lower thermodynamic stability at their 5' end are preferentially stabilized to perform their function in gene silencing (Khvorova et al. 2003; Schwarz et al. 2003). Different forms of double stranded RNAs generate
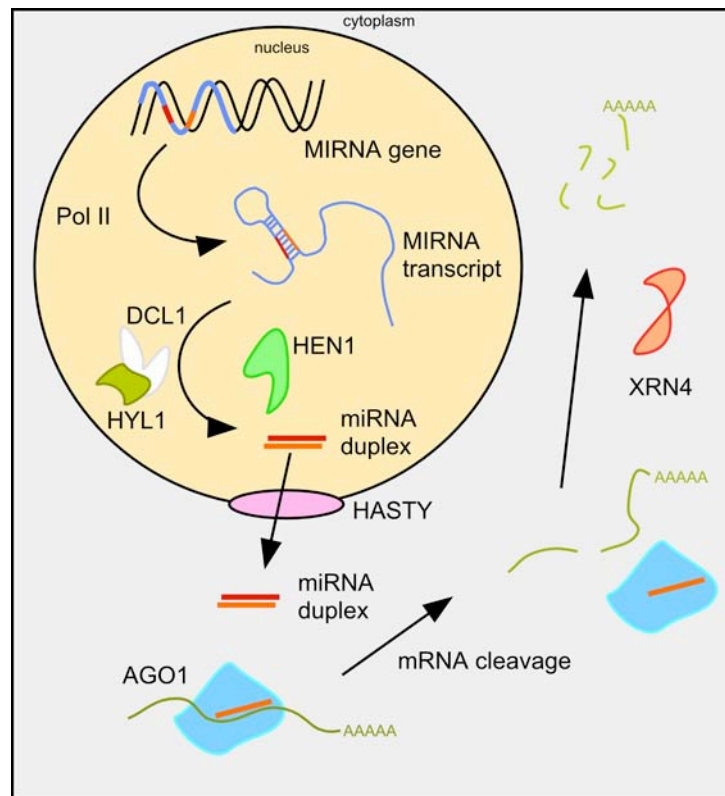
Figure 1: Biogenesis and action of microRNAs in plants
modified from Kidner et al. (2005)

different classes of small RNAs, such as siRNAs and miRNAs.

MiRNAs are encoded in the genome and primary transcripts, which are formed in the nucleus, are often several kilobases in length (Lee et al. 2002) see (Figure 1). They are transcribed mostly from independent transcription units with separate promoter elements. As they are generated by polymerase II (Lee et al. 2004), many promoter elements include TATA boxes and transcripts contain both 5' guanosine caps and 3' poly-adenosine tails, as do conventional protein-coding transcripts (Xie et al. 2005). MiRNA transcripts harbor one or several characteristic imperfect stemloops of ~70-80 nucleotides in animals, and ~80-250 nucleotides in plants with extensive, yet not perfect base-pairing in their stem region, which contains the future mature microRNA sequence. The same or very similar miRNA sequences are often produced from different related stemloop precursors, which together constitute a miRNA family, many of which are conserved among related species. Initially, miRNAs had been cloned from small RNA libraries, and more recently computational approaches have added several more candidates (see below). MiRNAs have been numbered according to their entry into the miRNA database rfam, which is located at http://microrna.sanger.ac.uk, and miRNAs of the same family but from different precursors are normally termed a, b, etc. The number

14

of miRNAs has been estimated to ~120 in *C.elegans* and ~250 in human, but as many approaches to predict miRNAs from genome sequences are based on conservation in closely related species (see below), some additional non-conserved members might have to be added to these numbers. Non-conserved miRNAs do exist in both plants and animals (Bentwich et al. 2005; Lu et al. 2005b), even though the great majority of known miRNAs can be found even in very distantly related species. There is, however, no common miRNAs in plants and animals, which might suggest that miRNAs have evolved independently in the two kingdoms. Plant miRNAs have been proposed to originate from inverted duplication of target gene sequences, followed by accumulation of mutations (Allen et al. 2004). Consistent with that idea, some non-conserved and thus potentially evolutionary young miRNAs are located adjacent to their respective target genes and contain sequence similarities to these genes also outside the small RNA sequence (Allen et al. 2004). Most miRNA target sequences in plants are found within coding regions of target genes, which is also in line with their proposed origin. Target sites of animal miRNAs on contrary are normally located in 3' UTRs of target genes and are often present in several copies. Additionally, binding sites for different miRNAs in one UTR are quite common, also indicating that animal miRNAs have evolved differently as their plant counterparts.

MiRNA precursors have also been indentified from pathogenic viruses (Cai et al. 2004; Grey et al. 2005; Pfeffer et al. 2005; Samols et al. 2005) and these most likely use the host machinery to produce active small RNAs.

In animals, the nuclear RNase Drosha recognizes the local double strand in miRNA precursors and cleaves off the stemloop from the remainder of the transcript, leaving a 2nt 3' overhang and a free 5' phosphate (Lee et al. 2003). The stemloop contains the future mature microRNA either in the 5' or 3' arm, and Drosha cleavage generates its 5' or 3' end. The stemloop is subsequently exported into the cytoplasm in a Ran-GTP dependent manner via the transport protein Exportin-5 (Bohnsack et al. 2004; Lund et al. 2004; Yi et al. 2003). In the cytoplasm, Dicer RNase processes the stemloop to form a ~19-23 nucleotide double stranded product, the miRNA/miRNA* duplex. In *Arabidopsis*, the nuclear Dicer family member DICER-LIKE 1 (DCL1) is thought to mediate both cleavage steps to form the miRNA containing short double strand, as has been shown for miRNA163 (Kurihara and Watanabe 2004). The *Arabidopsis* methyl-transferase HUA ENHANCER 1 (HEN1) has been shown to add methyl groups to the 3' terminal riboses of both strands in the miRNA/miRNA* duplex (Yu et al. 2005), which prevents addition of one or several uridyl residues to stabilize miRNAs (Li et al. 2005b). Similar observations have been made for siRNA double strands in plants, but not for small RNA intermediates in

animals. HASTY, the plant Exportin-5 homolog has been implicated in shuttling miRNAs into the cytoplasm, where they can exert their function (Bollman et al. 2003).

SiRNAs are typically produced from long, perfectly complementary precursors. These are formed for example during replication of viral RNA genomes and as a consequence of bidirectional transcription, which can occur in transposon- and repeat rich regions, but also when transgenes together with an exogenous promoter insert in close vicinity to an endogenous promoter, which transcribes in opposite direction. Furthermore, single stranded aberrant (mostly transgene or repeat) RNAs can be transformed into double stranded molecules with the help of RNA dependent RNA polymerases (see below). SiRNA precursors are processed by Dicer RNases (DICER-LIKE in plants) into several short double-stranded intermediates with unknown 5' ends.

Type III Ribonucleases

Type III Ribonucleases which process small RNA precursors in higher plants and animals are typically large enzymes and contain two RNase domains and a single dsRNA binding motif (Tomari and Zamore 2005). They split into two classes (Figure 2), the Drosha and the Dicer clade, with Dicer enzymes harboring additional functional domains. Among those a helicase moiety at their amino-terminal end and a PAZ domain, which is thought to specifically bind the single stranded tails of small RNA duplices. Drosha as well as Dicer RNases have been shown to mediate processing of siRNA or miRNA precursors, yet they do not function on their own, but require a double-stranded RNA-binding protein partner to mediate RNA cleavage at the required positions (Liu et al. 2003). R2D2, the binding partner of siRNA-producing Dicer-2 in *Drosophila melanogaster*, was shown to selectively bind to the siRNA/siRNA* double stranded end with higher thermodynamical stability, thus orienting the Dicer-2/R2D2 complex and enforcing stabilization of only the functional single stranded small RNA (guide strand) (Tomari et al. 2004), while the other strand (passenger strand) is being degraded. Similarly, protein partners have been identified in different species for other Dicer and Drosha members (Tomari and Zamore 2005) (Figure 2).

Whereas animals such as *C. elegans* or human contain only a single Dicer enzyme, generating both siRNAs and miRNAs, other species have split these functions between different proteins. *Drosophila* uses the Dicer-1 protein together with its binding partner Loquacious for miRNA processing (Forstemann et al. 2005; Saito et al. 2005) and Dicer-2/R2D2 to generate siRNAs. In *Arabidopsis*, 4 different

16

DICER-LIKE (DCL) proteins have been described with different, but also partially overlapping functions in small RNA biogenesis (Gasciolli et al. 2005). DCL1 is the main enzyme generating miRNAs, and mutants show pleiotropic phenotypes (Golden et al. 2002), consistent with miRNA function in development. HYL1, an R2D2 homolog has been shown to interact with DCL1 in vitro (Hiraguri et al. 2005), and *hyl1* mutants also show developmental defects (Vazquez et al. 2004a). DCL2 and DCL3 process viral and endogenous siRNAs respectively (Xie et al. 2004), while DCL4 has been implicated in the formation of secondary siRNAs, which are described in more detail below (Dunoyer et al. 2005).
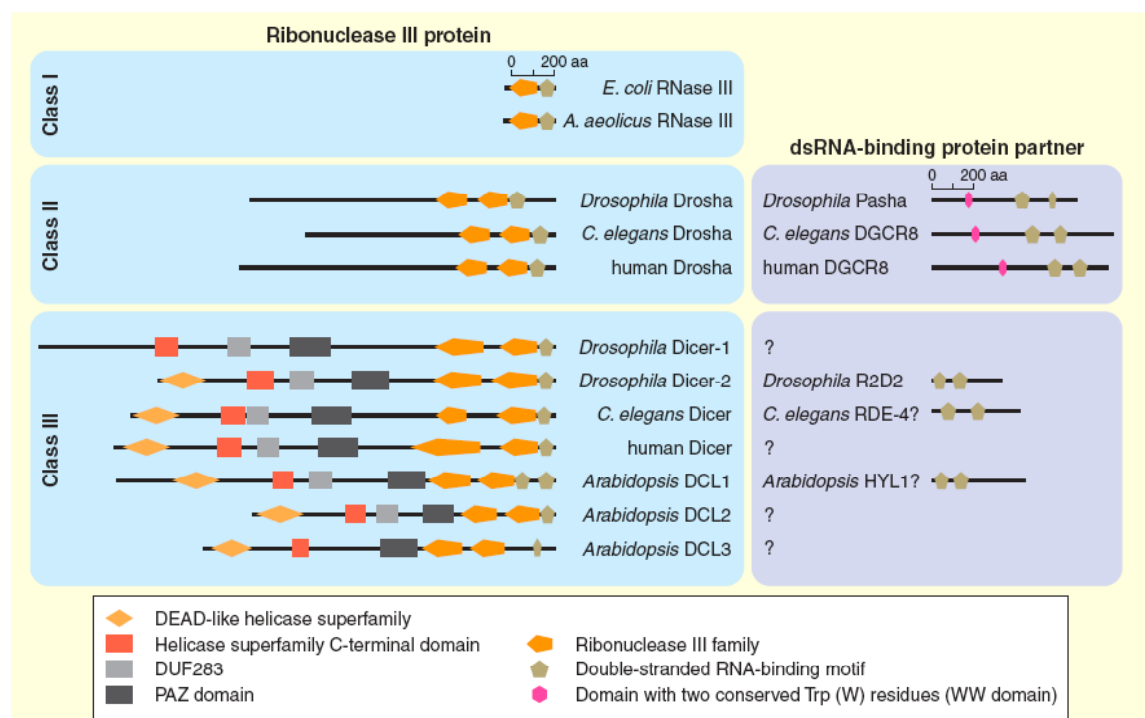
Figure 2: The family of RNaseIII enzymes involved in small RNA biogenesis
(Tomari and Zamore 2005)

Assembly of small RNA containing ribonucleoprotein complexes

SiRNAs serve as specificity components of the protein complex RISC (RNA induced silencing complex), which possesses endonucleolytic activity to cleave binding transcripts in the region pairing to the small RNA. Similarly, animal miRNAs have been found in protein complexes, generally termed miRNPs.

Argonaute proteins are the central players of RISC-like complexes and specifically bind small RNAs with their PAZ domains. The structure of the PAZ

domain has been solved by crystallization and NMR (Lingel et al. 2003; Song et al. 2003; Song et al. 2004; Yan et al. 2003) and it forms a binding pocket for ~21mer small RNAs, with a specific region recognizing the 2 nucleotide 3' overhang, which are typical for double stranded Dicer cleavage products. A second functional region of Argonaute proteins, the PIWI domain, has been implicated in target RNA cleavage ("slicer" activity), as shown for Argnonaute2 of *Drosophila melanogaster*, which processes siRNA target transcripts (Liu et al. 2004).

Similar to Dicer proteins, Argonaute proteins are not unique in many genomes, suggesting that they have acquired specialized functions. The *Arabidopsis* genome contains 10 Argonaute family members, whose functions are only partially understood. *AtAGO1* functions in both miRNA and siRNA target regulation and mutants show developmental abnormalities (Vaucheret et al. 2004). Unlike animal Argonaute proteins, which are normally found as part of a larger protein complex, AtAGO1 directs miRNA and siRNA target cleavage without requiring any protein partners (Baumberger and Baulcombe 2005). *AtAGO4* is involved in siRNA-dependent silencing of transposons and repeats (Zilberman et al. 2004), whereas *AtAGO7* (*ZIPPY*) has been implicated in mediating effects of endogenous, miRNA-dependent secondary siRNAs (trans-acting siRNAs) (Vazquez et al. 2004b), which influence phase change in developmental timing (Hunter et al. 2003) (see below). *AtAGO10* (*ZWILLE*) probably functions redundantly with *AtAGO1* since mutants also show developmental phenotypes (Lynn et al. 1999; Moussian et al. 1998).

RISC assembly subsequent to Dicer-mediated small RNA biogenesis has been studied extensively and recent experimental evidence supports the hypothesis, that Dicer, as Argonaute proteins, is parts of the RISC complex, so that Dicer cleavage products are directly handed over to the Argonaute PAZ domain, where passenger strands are selectively degraded. Orientation of the Dicer double stranded products, which require the dsRNA binding protein partners like R2D2 is based on thermodynamic characteristics (see above).

Mechanistic basis of small RNA-directed gene silencing

SiRNA function in post-transcriptional gene silencing normally becomes evident on the RNA level, as levels of target transcripts are reduced in the presence of siRNAs. Argonaute proteins direct cleavage of target transcripts opposite of position 10/11 of the small RNA, and exonucleases mediate degradation of the products.

Plant miRNAs function similarly to siRNAs, as they also trigger cleavage of their target transcripts (Kasschau et al. 2003; Llave et al. 2002b), and the

exonuclease *AtXRN4* has been implicated in destruction of the resulting RNA remnants (Souret et al. 2004). Consistently, *dcl1* mutants show increased accumulation of several miRNA target transcripts (Allen et al. 2004). However, feedback regulation of target transcription can interfere with full-length transcript abundance, postulating that in addition to target cleavage, plant miRNAs might also function as inhibitors of target translation (discussed in results). Another level of feedback regulation operates in plant systems, as both *DCL1* and *AGO1*, the main components of miRNA biogenesis and function, are themselves under miRNA regulation (Vaucheret et al. 2004; Xie et al. 2003).

Similar to *lin-4*, the first miRNA discovered in *C.elegans*, most animal miRNAs block translation of target genes, thereby reducing target protein abundance (Lai 2002). A recent report has suggested that *lin-4* targets are also reduced at their transcript levels, however, this mechanism relies on 5'->3' exonucleases and is not a direct miRNA effect (Bagga et al. 2005). Instead, decapping of target transcripts is likely to reduce their stability. Only few examples of animal miRNAs directly affecting transcript turnover (Jing et al. 2005; Yekta et al. 2004) or viral replication (Jopling et al. 2005) have been reported. The mechanistic basis of translational inhibition has been studied in less detail compared to transcript cleavage, but it has been reported that miRNA targets as well as Argonaute proteins associate to cytosolic P-bodies, which are general sites of mRNA degradation (Liu et al. 2005a; Liu et al. 2005b). As *lin-14*, a target of *lin-4* has been found to sediment in a polysomal fraction when applying a density gradient (Olsen and Ambros 1999), it has been assumed that inhibition of translation occurs after its initiation. However, a recent report suggests that translational inhibition indeed does occur at the level of its initiation, as both 5'cap and 3' poly-A tails of mRNA targets are necessary for miRNA mediated repression(Humphreys et al. 2005).

Complexity of small RNA pathways

(summarized in Figure 3)

In addition to directing cleavage of longer target RNAs, many siRNAs can also trigger formation of secondary siRNAs to amplify the silencing signal. This mechanism relies on RNA dependent RNA polymerases (RdRPs, RDRs), which extend local RNA double strands generated by siRNA binding to a target RNA, and form longer double stranded products (Maine 2000). This product can subsequently be processed by Dicer RNases to form secondary siRNAs, which may or may not be related in sequence to primary siRNAs. This process is generally termed transitivity (Himber et

al. 2003) and is not functional in all organisms, as RdRPs are absent from e.g. the *Drosophila* genome. In *C.elegans*, local double stranded RNA regions can be extended in the 5' direction, whereas, both directions are formally possible in plants. An additional feature of plant siRNAs is their ability to move across short distances (~10 cell layers) via plasmodesmata (Dunoyer et al. 2005). Formation of secondary siRNAs can therefore also spread the silencing trigger throughout the organism, as several rounds of amplification can involve movement of siRNAs prior to RdRP-dependent double strand formation. This process has been associated with plant resistance to viruses in non-infected tissues after local infection and before spreading of the virus itself (Voinnet et al. 1998). Replicating viral RNA genomes can serve as ideal Dicer substrates and are processed into small siRNAs. Some viruses, however, encode so-called silencing suppressor proteins, which interfere with small RNA pathways at different steps and prevent the silencing machinery from blocking viral spreading (Voinnet 2005).

In *Arabidopsis*, primary siRNA (~21nt size class) formation involves *DCL2* in case of virus and transgene derived substrates and *DCL3* for chromatin related silencing. *AGO1* has been implicated in target RNA cleavage (Vaucheret et al. 2004). Formation of secondary siRNAs uses *RDR6*, an RdRP, and *SDE3*, a helicase, to generate double stranded RNA, which serves as Dicer substrate (Dalmay et al. 2001). *DCL4* has been implicated in formation of 21nt siRNAs, which can move through plasmodesmata, but not in generation of 24-25 nucleotide siRNAs (Dunoyer et al. 2005).

Plant miRNAs, which cleave target RNAs similar to siRNAs, normally do not trigger formation of secondary small RNAs (Lu et al. 2005a). However, two cases have been reported, where miRNAs bind non-coding RNAs as primary targets, and direct both cleavage and secondary small RNA formation (Allen et al. 2005). These secondary RNAs are termed trans-acting siRNAs (tasiRNAs) and subsequently target a set of protein coding mRNAs, thus regulating e.g developmental phase change, which also involves *AGO7* (Vazquez et al. 2004b). The role of these miRNAs, miR173 and miR390, in secondary siRNA formation includes determination of the ~21nt phase, by which *DCL4* cleaves its double-stranded products (Allen et al. 2005). Unlike most other siRNAs, tasiRNAs perform their function in *cis* on RNAs they do not originate from.

Similar observations have been reported by Borsani and colleagues (Borsani et al. 2005), who identified the formation of a ~24nt small RNA from an antisense overlapping gene pair, setting the phase for additional formation of ~21nt small RNAs. These in turn function to silence one of the two transcripts they originate from.

A second class of siRNAs in plants (24-25 nucleotides) has been described as triggers of RNA-dependent DNA methylation (RdDM). This process describes methylation of cytosine residues in DNA regions, which are homologous to the small RNA sequence. It is often coupled to methylation of histones (mostly lysine 9 residues on histone H3) in the corresponding regions (Matzke et al. 2004). Small RNA directed histone methylation has also been described in fission yeast (*Schizosaccharomyces pombe*), however without modifications of DNA (Grewal and Rice 2004). Both DNA and histone methylation can result in heterochromatin formation, and covalent modifications of DNA or histones classifies them as epigenetic processes. As transcription is not involved in gene silencing by chromatin methylation, this process is termed transcriptional gene silencing (TGS) and occurs mainly in repeat and transposon rich regions. It involves components of the silencing machinery, such as the RNA dependent RNA polymerase *RDR2* and *DCL3* (Xie et al. 2004), and also Swi/Snf-related chromatin remodeling factors, such as *DRD1*, which is thought to regulate accessibility of DNA sequences for small RNA silencing triggers (Kanno et al. 2004). *AGO4* mediates small RNA effects to recruit methyltransferases, which transfer methyl groups on cytosines in CG islands, as well as CNG and CNN configurations (Zilberman et al. 2004). RNA dependent DNA methylation also involves the plant-specific polymerase IV at different steps with a hypothesized function in transcription of silenced transposons and repetitive elements (Herr et al. 2005; Kanno et al. 2005; Onodera et al. 2005).

Small RNA directed cytosine methylation has also been described for miRNA165/166, which targets HD-ZIP transcription factors such as *PHABULOSA* (*PHB*) and *PHAVOLUTA* (*PHV*). In this specific case, methylation occurs on the target transcripts and its biological relevance remains obscure (Bao et al. 2004).

Taken together, small RNA-directed gene silencing summarizes a variety of processes with highly specialized functions. All small RNAs are derived from double stranded RNA by means of cleavage by Dicer RNases. MiRNAs are mostly derived from endogenous hairpins and trigger post-transcriptional gene silencing (PTGS) in *trans* by means of target RNA cleavage in plants and translational inhibition in animals. They regulate mostly endogenous processes such as developmental timing. SiRNAs mostly direct PTGS in *cis*, to silence e.g. foreign RNA both locally and systemically. They are also involved in TGS, which directs heterochromatin formation, especially in repeat- and transposon rich regions.
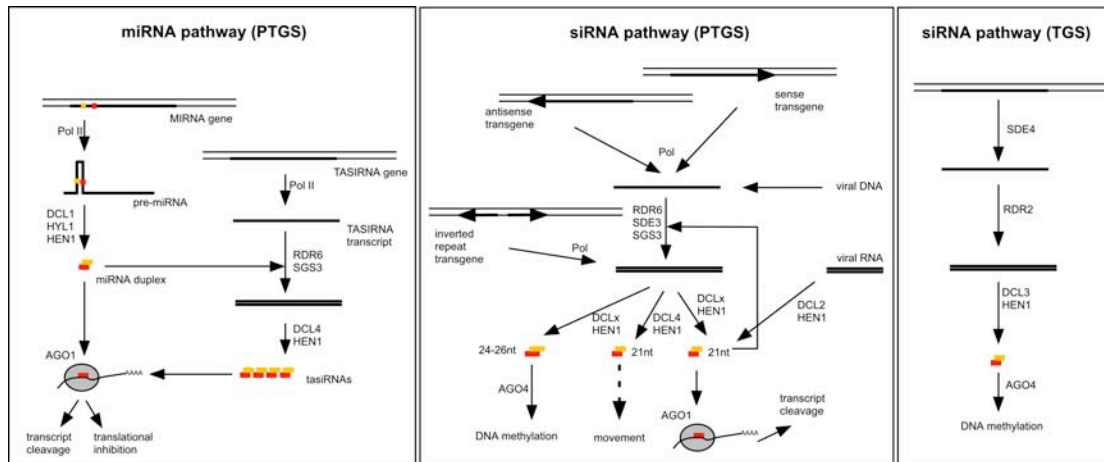
Figure 3: Diversity of small RNA pathways in plants

Specificity determinants of small RNA mediated gene silencing

SiRNAs mostly function in *cis* and silence RNAs they originate from. Therefore, most siRNAs share perfect match pairing to target RNAs. However, biochemical studies in human cell lines have demonstrated that pairing to small RNA positions 2 to 12 are sufficient to efficiently reduce target mRNA abundance (Doench and Sharp 2004). Similar experiments using *Drosophila* embryo lysates demonstrated that both terminal as well as central mismatches to target RNAs are tolerated by RISC with fewest mismatches tolerates in the 5' part of the small RNA (Haley and Zamore 2004). In line with these results, microarray analysis of human cells transfected with a siRNA demonstrate that not only perfectly complementary RNAs can be silenced by siRNAs (Jackson et al. 2003). However, consensus target determinants of siRNAs in animals have not yet been determined. Similarly, plant siRNA targeting has not been analyzed in detail.

Known plant miRNA targets share 0-5 mismatches with their miRNA and they function by target cleavage, similar to siRNAs (see Table 1 for a summary of plant miRNA families). Summarizing mismatches between all known targets and their miRNAs suggests that pairing to the 5' and central part of the small RNAs is most important, since mismatches are mostly found towards the 3' end (Mallory et al. 2004b). However, miRNA targets have mostly been proposed from computational predictions, which rely on low false positive rates and often on phylogenetic relation between targets of the same miRNA (Jones-Rhoades and Bartel 2004; Rhoades et al. 2002), which can be observed in many cases.

In animals, most known miRNA targets have many mismatches to their miRNAs and they are mostly regulated by translational inhibition. Computational efforts to predict new and additional targets in animals has resulted in high numbers of candidate target genes with low confidence for true positives (Enright et al. 2003; Kiriakidou et al. 2004; Lai 2002; Lewis et al. 2003; Stark et al. 2003). More recent studies in *Drosophila* and human have presented evidence that indeed pairing to only a small number of residues between miRNA and target transcript are sufficient for effective translational inhibition (Brennecke et al. 2005). The most important pairing region comprises nucleotides 2-8 of the miRNA and has been termed "seed region" (Brennecke et al. 2005; Lewis et al. 2005; Lim et al. 2005). Additionally, good pairing in the 3' part of the miRNA can compensate for mismatches in the seed region (Brennecke et al. 2005). Based on these studies, more than 100 targets for each animal miRNA have been suggested. In line with these suggestions, Lim and colleagues have shown that despite their major role in inhibiting translation, miRNAs cause slight transcriptional downregulation (probably indirect) of many transcripts, which share complementarity to the miRNA in its seed region. In support of the assumption that all these genes are really targeted by the miRNA, it was shown that transfection of undifferentiated cell cultures with a miRNA, which is normally preferentially expressed in a specific organ as e.g. liver, causes a shift of the transcriptional profile towards that of liver cells. In summary, animal miRNAs have many targets to which they share only limited complementarity and which are generally regulated by translational inhibition. Very high complementarity however can also cause cleavage of target transcripts as is normally observed in plants (Yekta et al. 2004).

Directed gene silencing by small RNAs

Small RNAs are effective regulators of gene expression and have been used to selectively silence genes of interests in different organisms. SiRNAs duplices can be directly applied to cell culture systems, such as human cell lines or *Drosophila* S2 cells, which are used for both genetic and biochemical analyses (Zamore et al. 2000). They are applied as *in vitro* synthesized RNA double strands with 2nt 3' overhangs and addition of different side chains, mostly 2' O-methyl groups, stabilizes the RNAs. In order to maximize effectiveness of siRNAs, many groups have optimized their sequence compositions by empirical testing of often several hundred different small RNAs. A comprehensive study by Reynolds et al (Patzel et al. 2005; Reynolds et al. 2004) describes sequence parameters, which are over-represented in

highly efficient siRNAs (>95% gene silencing). Among them are pronounced 5' instability, and an adenine nucleotide at position 10. Pairing of siRNA single strand terminal nucleotides has been shown to hinder silencing efficiency (Patzel et al. 2005). SiRNAs cannot only be transiently applied to cell culture systems, but they can also be stably integrated into host genomes by transformation with hairpin precursors, which generate siRNAs from appropriate promoter elements. Large libraries of so-called small hairpin (sh) RNAs have been generated containing precursors to individually silence most genes in the mouse and human genomes, which can be used for high throughput loss-of-function screens (Paddison et al. 2004). Similar approaches have been taken in *C.elegans*, where application of small RNAs is as simple as feeding worms on bacteria, which contain siRNA expressing plasmids (Timmons and Fire 1998).

In plants, siRNA directed gene silencing can be induced by expression of three different kinds of transgenes. High overexpression of sense-transgenes sometimes results in co-suppression of the respective transcript, so that very low, rather than high transcript levels are the result of over-expression, leading to loss-of-function effects (sense PTGS, s-PTGS). It is assumed that RNA-dependent RNA polymerases recognize the aberrant transgene RNA and direct double-strand formation, which induces Dicer-dependent formation of small RNAs and subsequent silencing of both the transgene and the endogenous gene. A higher frequency of gene silencing has been observed when expressing antisense transgenes from a strong promoter. It is assumed that sense and antisense strands hybridize to form a double-stranded product, which is in turn processed into small RNAs. The most efficient way to generate siRNAs in plants is from hairpin precursors, similar to animal systems (ir-PTGS). Here, 100-800 base-pair fragments of a gene of interest are cloned in sense and antisense direction between a loop-forming intron, so that primary transgene transcripts very efficiently form a perfect hairpin (Wesley et al. 2001). This hairpin is processed by Dicer to generate small RNAs, which direct gene silencing. This method has been widely used in *Arabidopsis* and other model plant species and often, but not always, resulted in efficient silencing of the gene of interest, which shares perfect complementarity to any small RNA derived from the hairpin precursor. The 5' ends of these small RNAs, however, are unknown as is their specificity of target selection. Therefore, it cannot be predicted, if transcripts other then of the gene of interest are affected by small RNA action.

Table 1 Plant microRNA families and their targets

miRNA	family members	predicted or confirmed target genes	proposed functions of targets	conservation in rice	Ref.
156/ 157	8/4	*SPL* transcription factors	flowering time regulation	+	1-3
158	2	xyloglucan fucosyl transferases, lipase		-	1, 2
159	3	*GAMYB* transcription factors	stamen development, flowering time regulation	+	1, 2, 4-6
160	3	Auxin response factors	auxin signaling	+	1, 2, 7, 8
161	1	*PPR* genes		-	1, 2, 9
162	2	*DCL1*	miRNA processing	+	1, 2, 10
163	1	methyltransferases		-	1, 2, 9, 11
164	3	*NAC* transcription factors	organ separation, meristem maintenance	+	1, 2, 3, 12-15
165/ 166	2/7	*HD-ZIP* transcription factors	organ polarity	+	1, 2, 16-21
167	4	Auxin response factors	auxin signaling	+	1, 2
168	2	*AGO1*	miRNA function	+	1, 2, 22
169	14	*HAP2* transcription factors		+	1, 2
170/ 171	1/3	*GRAS* transcripton factors	root patterning	+	1, 2, 23, 24
172	4	*AP2* transcription factors	flowering time regulation, floral patterning	+	5, 25-30
173	1	*TAS1,2*	generation of trans-acting siRNAs	-	25, 31, 32
319	3	*TCP* transcription factors	determining leaf shape	+	33
390	2	*TAS3*	generation of trans-acting siRNAs	+	31, 34
393	2	*TIR1*	auxin signaling	+	35, 36
394	2	F-box genes	protein degradation	+	35
395	6	ATP-sulfurylases		+	35
396	2	*GRL* transcription factors		+	35
397	2	Laccases		+	35, 36
398	3	Copper superoxid dismutases		+	35, 36
399	6	UBQ conjungating enzyme	protein degradation	+	35, 37, 38
400	1	*PPR* genes		-	36
401	1	unknown proteins		-	36
402	1	DNA glycosylase		-	36
403	1	*AGO2*	small RNA metabolism	-	31, 36
404	1	protein kinase		-	36
405	3	unknown protein		-	36
406	1	spliceosomal proteins		-	36
407	1	dehydrogenase		-	36
408	1	laccases and others		+	3,36

References:

1: (Reinhart and Bartel 2002), 2: (Rhoades et al. 2002), 3: (Schwab et al. 2005), 4: (Achard et al. 2004), 5: (Mette et al. 2002), 6: (Millar and Gubler 2005), 7: (Wang et al. 2005), 8: (Mallory et al. 2005), 9: (Allen et al. 2004), 10: (Xie et al. 2003), 11: (Kurihara and Watanabe 2004), 12: (Mallory et al. 2004a), 13: (Laufs et al. 2004), 14:

(Guo et al. 2005), 15: (Baker et al. 2005), 16: (Mallory et al. 2004b), 17: (Li et al. 2005a), 18: (Kidner and Martienssen 2004), 19: (Floyd and Bowman 2004), 20: (Williams et al. 2005), 21: (Kim et al. 2005), 22: (Vaucheret et al. 2004), 23: (Llave et al. 2002a), 24: (Parizotto et al. 2004), 25: (Park et al. 2002), 26: (Aukerman and Sakai 2003), 27: (Schmid et al. 2003), 29: (Lauter et al. 2005), 30: (Chen 2004), 31: (Allen et al. 2005), 32: (Yoshikawa et al. 2005), 33: (Palatnik et al. 2003), 34: (Adai et al. 2005), 35: (Jones-Rhoades and Bartel 2004), 36: (Sunkar and Zhu 2004), 37: (Fujii et al. 2005), 38: (Chiou et al. 2005).

Results

Many miRNAs have been cloned from small RNA libraries in plants and animals (Aravin et al. 2003; Arazi et al. 2005; Bentwich et al. 2005; Lagos-Quintana et al. 2003; Lagos-Quintana et al. 2002; Lu et al. 2005a; Lu et al. 2005b; Park et al. 2002; Reinhart et al. 2002; Suh et al. 2004; Sunkar et al. 2005; Sunkar and Zhu 2004), and additional numbers have been predicted by computational efforts based on the potential of their precursors to form characteristic secondary structures (Adai et al. 2005; Bonnet et al. 2004; Grad et al. 2003; Lai et al. 2003; Lee and Ambros 2001; Lim et al. 2003a; Lim et al. 2003b; Rehmsmeier et al. 2004; Wang et al. 2004b) and/or their conservation in related species (Berezikov et al. 2005; Dezulian et al. 2005; Li et al. 2005c).

Only few animal miRNA – target pairs have been characterized to date, and they show only partial sequence complementarity. Thus, prediction of additional targets for the same or other miRNAs with similar binding characteristics, postulated large quantities of animal miRNA targets with core binding requirements in the 5' part of miRNAs (Brennecke et al. 2005; Grun et al. 2005; Lewis et al. 2005; Lim et al. 2005).

For most plant miRNAs, putative target genes with high complementarity (0-4 mismatches) to the respective small RNAs have been identified computationally (Jones-Rhoades and Bartel 2004; Park et al. 2002; Rhoades et al. 2002), and cleavage products indicative of miRNA-mediated processing have been isolated for many of those (Allen et al. 2004; Aukerman and Sakai 2003; Kasschau et al. 2003; Llave et al. 2002b; Mallory et al. 2005; Mallory et al. 2004a; Palatnik et al. 2003; Park et al. 2002). Individual targets of one miRNA were often closely related family members and, in support of the authenticity of these genes as miRNA targets, the postulated sites of miRNA binding were often conserved in other plant species, whereas the surrounding nucleotides were not (Palatnik et al. 2003). Based on known physiological or developmental roles of these target genes, putative miRNA functions have been hypothesized and can be tested experimentally.

Whereas *lin-4*, the founder miRNA, had been identified as a loss-of-function mutant in a classical forward genetics screen (Lee et al. 1993), only few other examples of miRNA mutant alleles have been identified since (Baker et al. 2005; Guo et al. 2005). This is not only due to their small size, but also to redundancy between individual members of a miRNA family (Abbott et al. 2005). Genetic analysis of

miRNA function therefore requires tools to silence whole miRNA families or uses gain-of-function approaches. So-called antagomirs, small, stabilized RNAs of antisense complementarity to miRNAs, have been successfully used in mice to reduce levels of whole miRNA families and study loss-of-function phenotypes (Krutzfeldt et al. 2005). In plants, similar methodologies have not be been established to date, so that most functional studies use gain-of-function approaches (Achard et al. 2004; Aukerman and Sakai 2003; Chen 2004; Guo et al. 2005; Kim et al. 2005; Laufs et al. 2004; Mallory et al. 2005; Mallory et al. 2004a; Mallory et al. 2004b; Parizotto et al. 2004; Vaucheret et al. 2004; Wang et al. 2005). In addition, the role of miRNA mediated regulation of target RNAs is also analyzed by expression of miRNA-resistant versions of target transcripts, which are not altered in their amino acid content (Palatnik et al. 2003). Such dominant alleles have also been described from classical forward genetic screens (McConnell et al. 2001), emphasizing the importance of miRNA mediated gene regulation during plant development.

I sought to study the roles of three *Arabidopsis* miRNA families with potential roles in specification of floral organ identity and morphology as well as during the timing of the floral transition (Park et al. 2002; Rhoades et al. 2002). Known details of these miRNA families – miR156, miR164 and miR172 – and their target genes are given below.

CHAPTER I

CHARACTERIZATION OF ENDOGENOUS microRNA

FUNCTIONS IN *Arabidopsis thaliana*

Specific Effects of MicroRNAs on the Plant Transcriptome

Rebecca Schwab*, Javier F. Palatnik*, Markus Riester, Carla Schommer, Markus

Schmid and Detlef Weigel

* These authors contributed equally

Contributions to this chapter:

All experiments and their analysis described in this chapter have been carried out by myself, if not mentioned otherwise.

MicroRNA families and predicted target genes

MicroRNA156 and *SPL* genes

Arabidopsis miRNA156 is encoded by eight precursors, *MIR156a-h* (http://microRNA.sanger.ac.uk). They produce small RNAs with slight differences in sequence and size, the most abundant being a 20mer generated from precursors a-f, which is closely related to another miRNA family, miRNA157, with 4 members (Reinhart et al. 2002). Mature miRNA sequences are listed in Table 2. Expression of miR156/157 is predominant in vegetative tissues of *Arabidopsis*, and can also be detected in gymnosperms, ferns, lycopods and mosses, suggesting a long evolutionary history (Arazi et al. 2005; Axtell and Bartel 2005).

Based on sequence similarities between small RNA and putative miRNA binding elements (target sequences), Rhoades et al. (2002) have predicted 11 members of the *Arabidopsis SQUAMOSA PROMOTER BINDING PROTEIN LIKE* (*SPL*) family as target of miR156/157. The *SPL* family encodes 16 plant specific transcription factors (Cardon et al. 1999), named after their *Antirrhinum* homolog *SQUAMOSA PROMOTER BINDING PROTEIN 1* (*SBP1*), which binds to elements in the *SQUAMOSA* (*SQUA*) promoter in vitro (Klein et al. 1996). *SQUA*, as its *Arabidopsis* ortholog *APETALA1* (*AP1*), functions in specification of floral identity (Mandel et al. 1992). The *Arabidopsis* SPL3 protein can also bind the *AP1* promoter at a TCCGTACAA core element, however *AP1* function is not required for SPL3-mediated effects (Cardon et al. 1997). SPL proteins are diverse in sequence outside the common SBP domain, and only two members, both of which do not contain a miR156 target site, have been functionally characterized. SPL8 is involved in pollen development (Unte et al. 2003), while SPL14 mediates programmed cell death in response to the fungal toxin FB1 (Stone et al. 2005). However, for most *SPL* genes, which do contains a miR156 binding element, mutants have not been described. Whereas most *SPL* target genes contain the miRNA target sequence in their coding regions, it is located in the 3'UTR for *SPL3*, *4*, and *5*, which are the smallest family members. Overexpression of SPL3 cDNA, which does not contain the miRNA target site, has been shown to cause early flowering in transgenic plants (Cardon et al. 1997), suggesting a role of miR156 in negative regulation of flowering time. *SPL* target genes are listed in Table 3.

MicroRNA164 and *NAC* genes

The miR164 family comprises three members in *Arabidopsis* (*MIR164a-c*), encoding 2 different isoforms of miR164, which are listed in Table 2. Expression analysis has shown that miR164 is present in many tissues and developmental stages (Axtell and Bartel 2005; Mallory et al. 2004a) and it is conserved in other angiosperms and gymnosperms (Axtell and Bartel 2005).

Predicted targets of miR164 belong to the plant specific *NAC* transcription factor family, named after its founding members *NAM* (*NO APICAL MERISTEM)* in petunia, *ATAF* and *CUC* (*CUP SHAPED COTYLEDONS)* in *Arabidopsis* (Aida et al. 1997; Souer et al. 1996). *NAM* and *CUC* genes are involved in separation of cotyledons and floral organs and maintenance of the shoot apical meristem (SAM), while mutants in the phylogenetically closely related *CUP* (*CUPULIFORMIS*) gene from *Antirrhinum* maintain a functional SAM, but show defects in separation of lateral organs, resulting in fused leaves and also fasciated inflorescences (Weir et al. 2004). Other members of the large *Arabidopsis NAC* gene family have diverse functions, and *NAC1*, which has also been predicted as a miR164 target, functions as a mediator of auxin signaling in the root, regulating the initiation of lateral root primordia (Xie et al. 2000). Functional roles of other predicted targets have not been reported. They all carry the miRNA target site downstream of the conserved *NAC* DNA binding domain region, as do petunia *NAM* and *Antirrhinum CUP* (Mallory et al. 2004a).

MicroRNA172 targets AP2-like transcription factors

MiRNA172 is encoded by five precursors, *MIR172a-e*, which produce three different isoforms of miR172, differing only at their terminal nucleotides (see Table 2). Like miRNA156, miRNA172 is not specific to *Arabidopsis*, but can also be found in monocots, gymnosperms and ferns (Aukerman and Sakai 2003; Axtell and Bartel 2005). Expression analyses suggest that miR172 is increasingly accumulating when plants mature and most abundant in floral tissue, more specifically in young floral buds and later in the inner whorls of young flowers (Aukerman and Sakai 2003; Axtell and Bartel 2005; Chen 2004; Schmid et al. 2003).

MiRNA172 targets belong to the large plant specific family of *AP2*-like transcription factors (Riechmann and Meyerowitz 1998) and comprise the family founding member *APETALA2* (*AP2*), a floral homeotic gene, specifying sepal and petal identity in the outer whorls of the *Arabidopsis* flower. In accordance with the ABC model of floral organ specification, *ap2* mutant flowers lack petals and show carpeloid sepals (Bowman et al. 1989). Other miR172 targets have been described

as repressors of flowering time. While *toe1* mutants flower earlier than wild-type in inductive long-day conditions, and mutations in *TOE2* enhance this effect (Aukerman and Sakai 2003), overexpression of *SMZ* and *SNZ* causes late flowering (Schmid et al. 2003). Functional roles of *TOE3*, the last predicted miR172 target, have not been reported previously. Apart from *TOE3*, miR172 target genes are more strongly expressed during the vegetative phase of the plant life cycle, contrasting with miR172 itself (Aukerman and Sakai 2003; Schmid et al. 2003).

Table 2: Genomic locations and mature sequences of microRNAs 156/157, 164 and 172

MIRNA	old name	next upstream gene	next downstream gene	mature sequence 5'->3'
156a		At2g25090	At2g25100	UGACAGAAGAGAGUGAGCAC
156b		At4g30970	At4g30980	UGACAGAAGAGAGUGAGCAC
156c		At4g31875	At4g31880	UGACAGAAGAGAGUGAGCAC
156d		At5g10940	At5g10950	UGACAGAAGAGAGUGAGCAC
156e		At5g11970	At5g11980	UGACAGAAGAGAGUGAGCAC
156f		At5g26140	At5g26150	UGACAGAAGAGAGUGAGCAC
156g		At2g19410	At2g19430	CGACAGAAGAGAGUGAGCACA
156h		At5g55830	At5g55840	UUGACAGAAGAAAGAGAGCAC
157a		At1g66780	At1g66790	UUGACAGAAGAUAGAGAGCAC
157b		At1g66790	At1g66800	UUGACAGAAGAUAGAGAGCAC
157c		At3g18215	At3g18220	UUGACAGAAGAUAGAGAGCAC
157d		At1g48740	At1g48750	UGACAGAAGAUAGAGAGCAC
164a		At2g47580	At2g47590	UGGAGAAGCAGGGCACGUGCA
164b		At5g01740	At5g01750	UGGAGAAGCAGGGCACGUGCA
164c		At5g27800	At5g27810	UGGAGAAGCAGGGCACGUGCG
172a	172a1	At2g28050	At2g28060	AGAAUCUUGAUGAUGCUGCAU
172b	172a2	At5g04270	At5g04280	AGAAUCUUGAUGAUGCUGCAU
172c	172b1	At3g11430	At3g11440	AGAAUCUUGAUGAUGCUGCAG
172d	172b2	At3g55510	At3g55520	AGAAUCUUGAUGAUGCUGCAG
172e	172c	At5g59500	At5g59510	GGAAUCUUGAUGAUGCUGCAU

Table 3: MicroRNA complementary motifs in predicted target genes of microRNAs 156/157, 164 and 172

Gene name	Identifier	Target sequence 5'->3' (mismatches to miR156a/164a/172a in bold)	Number of mismatches
SPL2	At5g43270	GUG CUC **U**CU CUC UUC UGU CA	1
SPL3	At2g33810	**U**UG CU**U** ACU CUC UUC UGU CA	2
SPL4	At1g53160	**C**UG CUC **U**CU CUC UUC UGU CA	2
SPL5	At3g15270	**CC**G CUC **U**CU CUC UUC UGU CA	3
SPL6	At1g69170	GUG CUC **U**CU CUC UUC UGU CA	1
SPL9	At2g42200	GUG CUC **U**CU CUC UUC UGU CA	1
SPL10	At1g27370	GUG CUC **U**CU CUC UUC UGU CA	1
SPL11	At1g27360	GUG CUC **U**CU CUC UUC UGU CA	1
SPL13	At5g50570 At5g50670	GUG CUC **U**CU CUC UUC UGU CA	1
SPL15	At3g57920	GUG CUC **U**CU CUC UUC UGU CA	1
CUC1	At3g15170	**A**GC ACG UG**U** CCU G**U**U UCU CCA	3
CUC2	At5g53959	**A**GC ACG UG**U** CCU G**U**U UCU CCA	3
NAC1	At1g56010	**A**GC ACG U**AC** CCU GCU UCU CCA	2
ANAC079	At5g07680	**U**U**U** ACG UGC CCU GCU UCU CCA	2
ANAC100	At5g61430	**UC**U ACG UGC CCU GCU UCU CCA	2
ANAC092	At5g39610	**CU**C ACG UG**A** CCU GCU UCU CC**G**	4
AP2	At4g39620	**C**UG CAG CAU CAU CA**G** GAU UCU	2
TOE1	At2g28550	**CA**G CAG CAU CAU CA**G** GAU UCU	3
TOE2	At5g60120	AUG CAG CAU CAU CA**G** GAU UCU	1
TOE3	At5g67180	**UG**G CAG CAU CAU CA**G** GAU UCU	3
SMZ	At3g45990	**U**UG CAG CAU CAU CA**G** GAU UC**C**	3
SNZ	At2g39250	**U**UG **U**AG CAU CAU CA**G** GAU UC**C**	4

Mapping of *MIRNA* transcription starts by RACE-PCR

This work was done in collaboration with Markus Schmid

MiRNAs are produced from fold-back precursors, which are normally part of a longer primary transcript, called pri-miRNA in animals. Two *Arabidopsis* EST clones encoding *MIR156c* and *MIR172a* precursors (accession numbers AK117457 and AK118705) have been sequenced and suggest that primary miRNA transcripts in plants contain exon-intron structures and are several hundred basepairs in length even after splicing events. As Polymerase II transcripts, they also contain 5' cap structures and poly-A tails (Lee et al. 2004).

In order to determine transcript structures of *MIRNA*156, 164 and 172 family members, we carried out 5' and 3' RACE-PCRs using a set of nested primer pairs located in and next to the stemloop sequences (oligo sequences are listed in Materials and Methods). RACE libraries were generated from floral RNA using a commercially available kit (SMART^TM RACE, Clontech).

We could amplify transcript fragments for several miRNAs, most of them determining transcript 5' ends. These were of different lengths and the miRNA containing stemloop was always located in the first exon. Interestingly, *MIR164* transcripts started very close to the stemloop forming sequences (34 nt for *MIR164a*, 85 for *MIR164b*), while much longer RACE fragments (239 to 506nt) were amplified for *MIR156* and *MIR172* family members (391 to 611nt). All amplified products are listed in Table 4.

A recently published report by Xie et al. (Xie et al. 2005) documents similar transcript starts for most *MIRNAs*, many of which are preceded by TATA-boxes at around -30 nucleotides. We used the information from 5' RACEs to determine upstream sequences of potential regulatory function specifying expression patterns of *MIRNA* genes.

Table 4: Mapping of *MIRNA* transcript starts by RACE-PCR

MIRNA	distance to next upstream gene	length of longest 5'RACE product relative to stemloop	5'RACE products determined by Xie et al. (2005)
*MIR*156a	3212 bp	506 bp	506 and 320 bp
*MIR*156b	310 bp	239 bp	not determined
*MIR*156c	3923 bp	410 bp (362 from EST)	325 and 168 bp
*MIR*156d	2807 bp	425 bp	not determined
*MIR*164a	2065 bp	34 bp	350 and 34 bp
*MIR*164b	6302 bp	85 bp	83 bp
*MIR*172a	7144 bp	not determined	584 bp (574 from EST)
*MIR*172b	4159 bp	611 bp	616 bp
*MIR*172e	1252 bp	391 bp	397 bp

Expression patterns of *MIRNA* promoter fragments

Spatial and temporal distribution of miRNAs in *Arabidopsis* has been studied by small RNA northern blotting (Aukerman and Sakai 2003; Kasschau et al. 2003; Mallory et al. 2004a; Palatnik et al. 2003; Reinhart et al. 2002), hybridization of different tissue types to microRNA-microarrays (Axtell and Bartel 2005) and miRNA *in situ* hybridizations (Chen 2004; Kidner and Martienssen 2004). All three methods mirror expression of miRNA families as they rely on hybridization of miRNAs to synthetic miRNA antisense probes, which cannot distinguish between sequence variants, which are typical for miRNA family members.

In order to determine expression domains of individual *MIRNA* genes, and analyze the degree of spatial redundancy between the the family members, I fused potential 5' regulatory regions of *MIRNA* genes, located upstream of the determined transcription starts, to the reporter gene *GUS* (Jefferson 1989). A similar approach has been taken by Parizotto and colleagues (Parizotto et al. 2004), who published their findings while my work was in still in progress. The authors have determined expression of *MIR*171 by fusion of 5' regulatory elements (1.2kb upstream of the stemloop) to the *GFP* reporter gene. They confirmed that the observed patterns corresponded to regions of miR171 activity by generating a miR171-reponsive *GFP* sensor transgene, which is only active in cells, which do not contain mature miR171 (the miR171 family consists only of a single member). The almost perfect complementarity of *GFP* activity when comparing promoter activity and sensor degradation also suggested that miRNAs do not move from cell to cell (Parizotto et al. 2004).

Table 5 shows the lengths of MIRNA promoter fragments, which were amplified from Col-0 genomic DNA and placed in front of the *GUS* open reading frame in pRITA1 (see supplementary material). The promoter – *GUS* – 3' nos terminator fragments were transferred to the binary plasmid pART27 (see supplementary material), which confers kanamycin resistance to transgenic plants.

For miR156c, I also extended the promoter fragment in the 3' direction, including additional 261bp of transcribed sequence preceding the stemloop precursor (RS382), to find out if this region contains additional information determining spatial or temporal expression.

Lines of representative patterns of *GUS* expression (all transgenes but RS382) were propagated to the T4 generation to carry out detailed histological analyses. Incubation in X-Gluc containing staining solution was carried out overnight.

Table 5: *MIRNA* promoter fragments

miRNA locus	transcript length 5' of stemloop	promoter length; transcribed part in parentheses	distance from stemloop to next upstream gene	binary plasmid name
*MIR*156c	362 bp	2377 (102) bp	3923 bp	RS159
*MIR*156d	425 bp	2404 (20) bp	2807 bp	RS205
*MIR*172a	574 bp	2204 (34) bp	7144 bp	RS157
*MIR*172b	611 bp	2399 (38) bp	4159 bp	RS158
*MIR*172e	391bp	858 (34) bp	1251 bp	RS206

***MIR156*.** Very strong and uniform activity of the *GUS* reporter gene driven by putative promoter fragments of *MIR*156c (both RS159 and 382) was observed in first true leaves of young long-day grown seedlings (Figure 4B-C). For *MIR*156d, staining was refined to the apical region surrounding the meristem and to the vasculature of cotyledons (Figure 4A). In rosette leaves, expression was most strongly observed in the vasculature in case of *MIR*156d (RS205; Figure 4D), while weaker uniform staining was detected with the shorter *MIR*156c (RS159; Figure 4E) fragment. Addition of the 261bp transcribed sequence of *MIR*156c in RS382 again refined *GUS* expression to the leaf vasculature (Figure 4F). Strongest expression in older seedlings was observed in the region around the shoot apex, where new leaves emerge, for all three transgenes (Figure 4G-I). In floral tissue, activity of the putative *MIR*156c promoter was observed in sepals, anthers and pollen of older, but not in younger flowers (Figure 4K-M). While the shorter RS159 fragment produced rather uniform staining in sepals (Figure 4L), the longer RS382 fragment refined expression to the vasculature (Figure 4M), as has also been observed in leaves. Additional staining in the style region of older carpels was only observed with RS159, but not with RS382.

Expression of *SPL* target genes is downregulated in stamens and pollen, consistent with high expression of *MIR*156 in this region.

***MIR172*.** Putative promoter elements of *MIR*172e (RS206) only activated the *GUS* reporter gene in the root vasculature in both vegetative and reproductive stages (Figure 5A, and not shown). *pMIR172a* (RS157) showed similar activities in roots, and also additional staining in young leaves (Figure 5A) and young floral organs (Figure 5B). Strongest staining was observed in young developing petals. Putative regulatory elements of *MIR172b* were only active during reproductive stages and very strong staining was observed in the inflorescence stems, and also in older

36
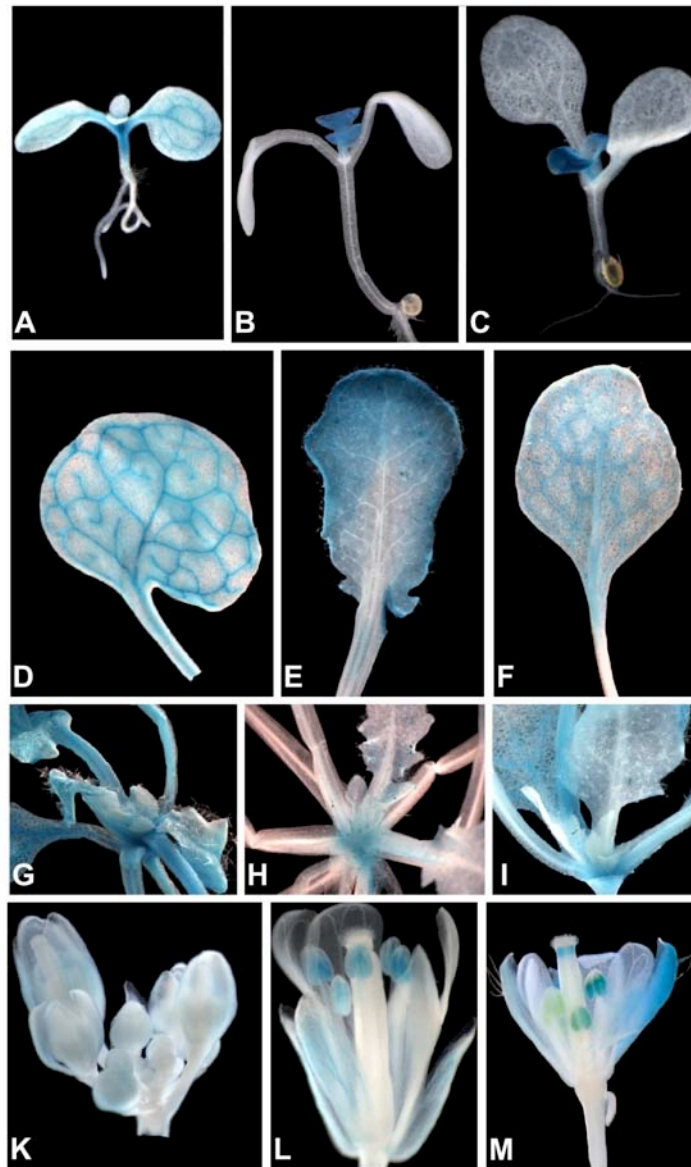
Figure 4: Patterns of *MIR156* promoter expression

(A-C) Young seedlings. Strongest *GUS* activity is apparent in the shoot apex (*pMIR156d* , RS205) **(A)** or young leaves (*pMIR156c* (RS159, RS205) **(B, C). (D-F)** Patterns of *GUS* expression in representative short-day grown rosette leaves. **(D)** *pMIR156d* (RS205), **(E)** short *pMIR156c* (RS159), **(F)** long *pMIR156c* (RS382). Note vein-restricted pattern in RS382, which is absent in RS159. **(G-I)** Older short-day grown seedling with strongest expression around the shoot meristem. **(G)** *pMIR156d* (RS205), **(H)** short *pMIR156c* (RS159), **(I)** long *pMIR156c* (RS382). **(K-M)** floral tissue. **(K)** No staining is detected in young inflorescences staining for *pMIR156c* (long fragment) activity. **(L)** Older flowers of the same genotype show strong staining in sepal veins and mature anthers. **(M)** The short *MIR156c* promoter fragment shows additional expression in the style, but signals in sepals are not restricted to the vasculature.

carpels (Figure 5B).

Since *AP2* activity is required for specification of sepal and petal identity in the outer whorls of the flower, the activity of the putative *MIR172a* promoter was not expected in these whorls. Therefore, closer inspections of the promoter elements, which have been analyzed, will be required to judge the authenticity of *MIR172* expression patterns. In this course, the long 5' ends of *MIR172* transcripts preceding the miRNA producing stemloops might play additional roles.
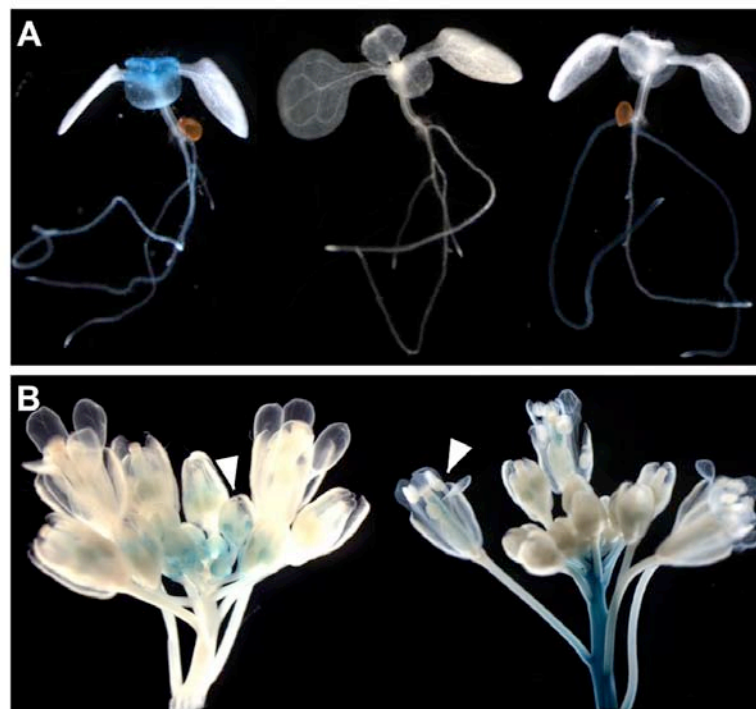
Figure 5: Staining patterns of *MIR172* promoter-*GUS* fusions

(**A**) Young seedlings. Left *pMIR172a*, middle *pMIR72b*, right *pMIR172e*. (**B**) Inflorescences. Left *pMIR172a*. Note staining in petals of young flowers (arrowhead). Right *pMIR172b*. Note staining in older carpels (arrowhead).

Overexpression of microRNAs in *Arabidopsis*

Plant miRNAs typically regulate a small number of target transcripts by simultaneous negative interference with their stability. To uncover the biological impact of such regulations, the system can be modified by changing the levels or expression domains of individual components. As miRNA mutants are rare and not easy to generate in plants, increased and ectopic negative regulation of target genes by

ubiquitous and constitutive expression of miRNA precursors is widespread to study miRNA functions (see references above).

Using the 35S promoter from cauliflower mosaic virus (CaMV35S), I overexpressed precursors from miRNA families 156, 164, and 172, and analyzed morphological changes in transgenic plants. Using several different precursors from each family, allowed me to test if all precursors were functional and could produce active miRNAs. To define, if sequences limited to the foldback structures were sufficient for miRNA production by *DCL1* or if outside regions contained necessary regulatory elements, I overexpressed sequences of different length, ranging from ~130bp limited to the foldback structures to >3kb genomic regions surrounding miRNA containing stemloops, thereby including most of their transcribed regions.

Overexpression of miR156 increases biomass and the rate of rosette leaf initiation

Arabidopsis miRNA156 has been predicted to regulate *SPL* transcription factors, among them *SPL3*, which confers early flowering to transgenic plants when overexpressed without the miRNA target site, which is located in its 3'UTR (Cardon et al. 1997). This finding suggests that miRNA156 might be involved in flowering time regulation, more precisely function as a floral repressor.

I amplified genomic fragments containing MIRNAs156a, b, c, d, and f from Col-0 genomic DNA as summarized in Table 6 and placed them between the CaMV35S promoter and the 3'OCS terminator in the pBJ36 derivative pMS37 (see Materials and Methods). The complete expression cassette was shuttled to the binary plasmid pMLBart (see Materials and Methods), which confers resistance to gluphosinate ammonium (Basta) in transgenic plants.

Table 6: MicroRNA156 overexpressing constructs

binary plasmid	*MIRNA*	size of transgene	sense oligo 5'->3'	antisense oligo 5'->3'
RS104	*MIR156a*	2272 bp	gattaggtgcctacatatac	gttcaccaatattccatgtcttc
RS105	*MIR156b*	1696 bp	gtaagacacgtgtagaaatc	cttcagggtgaagcacattag
RS107	*MIR156d*	2242 bp	ctcgttacccaaaatgaac	gggagggagaattctcaatttg
RS109	*MIR156f*	2794 bp	ggattcgtggtatagtgttac	ggctcatgttggaattcgaatc
RS116	*MIR156c*	135 bp	gacaaattttaagagaaacgcatag	gggaccgaatcggagccggaatctg

Transgenic plants were grown in both non-inductive short-day (SD) and inductive long-day (LD) conditions and flowering time was recorded as both the days to flower and total leaf number (TLN), which are normally tightly linked. As expected, miR156 overexpressers flowered slightly late after ~25 days in long days (averaged from 39 T1 plants overexpressing miR156b), whereas Col-0 wild-type flowered after ~19 days. All five transgenes caused more or less similar phenotypes, with slight variation in severity, whereas a population of primary transformants carrying the same transgene appeared rather uniform and phenotypic penetrance was close to 100% for all transgenes. These findings suggest that all precursors are capable of producing functional small RNAs and that sequences outside the fold-back structure are not required for *DCL1* action, since plants carrying the RS116 transgene of only 135bp showed similar defects as plants with longer miRNA156 precursor transgenes. If not otherwise mentioned, miR156b overexpressers (RS105) were used for further phenotypic characterization.

In addition to a slight temporal delay of flower production, miR156 overexpressers initiated rosette leaves notably faster than wild-type plants, so that total leaf number was greatly increased when flowers were first produced (Figure 6 A). Figure 8 A and B show the rate of leaf initiation, generally termed plastochron, for all five lines in SD and LD conditions.

Increased miR156 levels (see Figure 7 for a small RNA northern blot) also caused a severe decrease of apical dominance, such that first flowers often arose from side shoots and plant height was reduced. Combination of these traits led to a considerable increase in both side shoot number (especially higher order shoots; Figure 8C and 6B, D) and absolute number of leaves (Figure 8D and 6C, D), which could be more than ten times higher than in wild type, giving adult plants a very bushy appearance

Both fresh and dry weight of adult plants overexpressing miR156 was significantly higher than wild-type. However, at the time of wild-type maturity (~37days in LD), miR156 overxpressing plants were still growing and only initiated their reproductive phase, so that at both fresh and dry weight was very similar at that time (Figure 8 E, F).

When transforming the ecotype Landsberg *erecta* (L*er*) with the same *MIR*156b plasmid, primary transformants showed similar defects as described above for the ecotype Columbia (Col-0), but were much less drastic in long-day conditions (short days not tested).

Expression of *MIR*156b from the epidermal *ML1* promoter (Sessions et al. 1999) also increased both plastochron and side shoot number in transgenic plants

40

(RS248), but both features were much less pronounced as with the strong 35S promoter (not shown).
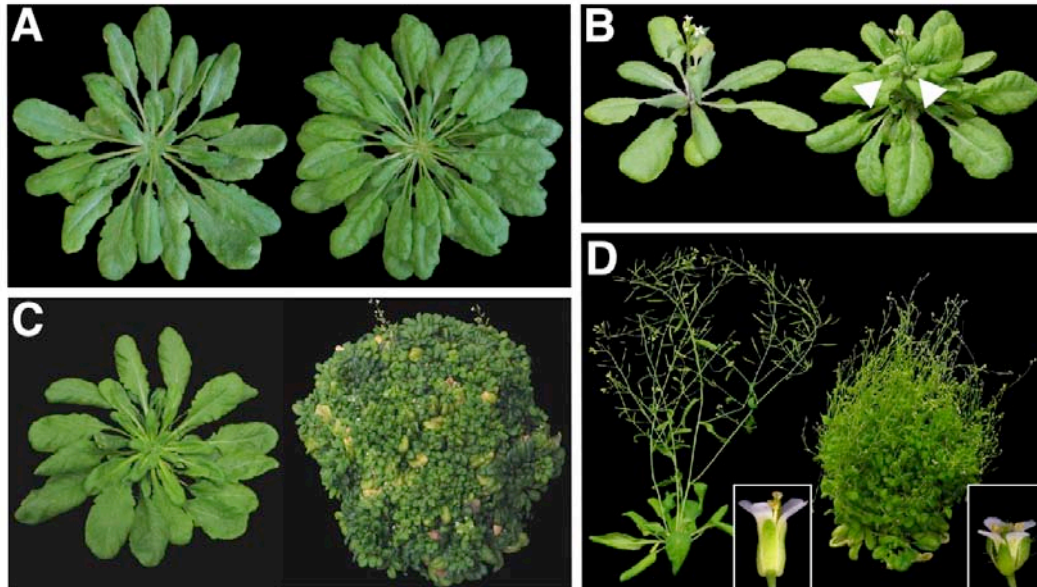
Figure 6: Phenotypes of plants overexpressing miRNA156b

(**A**) Plants grown in short days for 56 days. Note increased leaf number in *35S:miR156b*. (**B**) Plants grown in long days, shown shortly after main inflorescence has started to elongate. Precocious release of side shoots (arrow heads) is apparent in *35S:miR156b*. (**C**) Short-day grown plants shortly after the first open flowers have become visible. The *35S:miR156b* plant, which is about seven months old, has many more, but smaller leaves. (**D**) Long-day grown, mature plants at the stage that fruits are fully developed. Insets show flowers, highlighting squashed appearance of *35S:miR156b* flower.

.
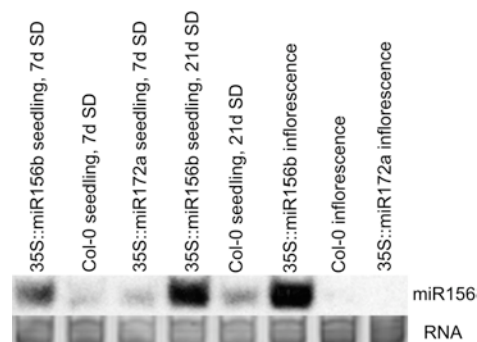
Figure 7: Overexpression of miRNA156 at different plant ages

Top panel: Small RNA northern blot. miRNA156 was detected with an end-labeled antisense DNA oligo. Bottom panel: Loading control
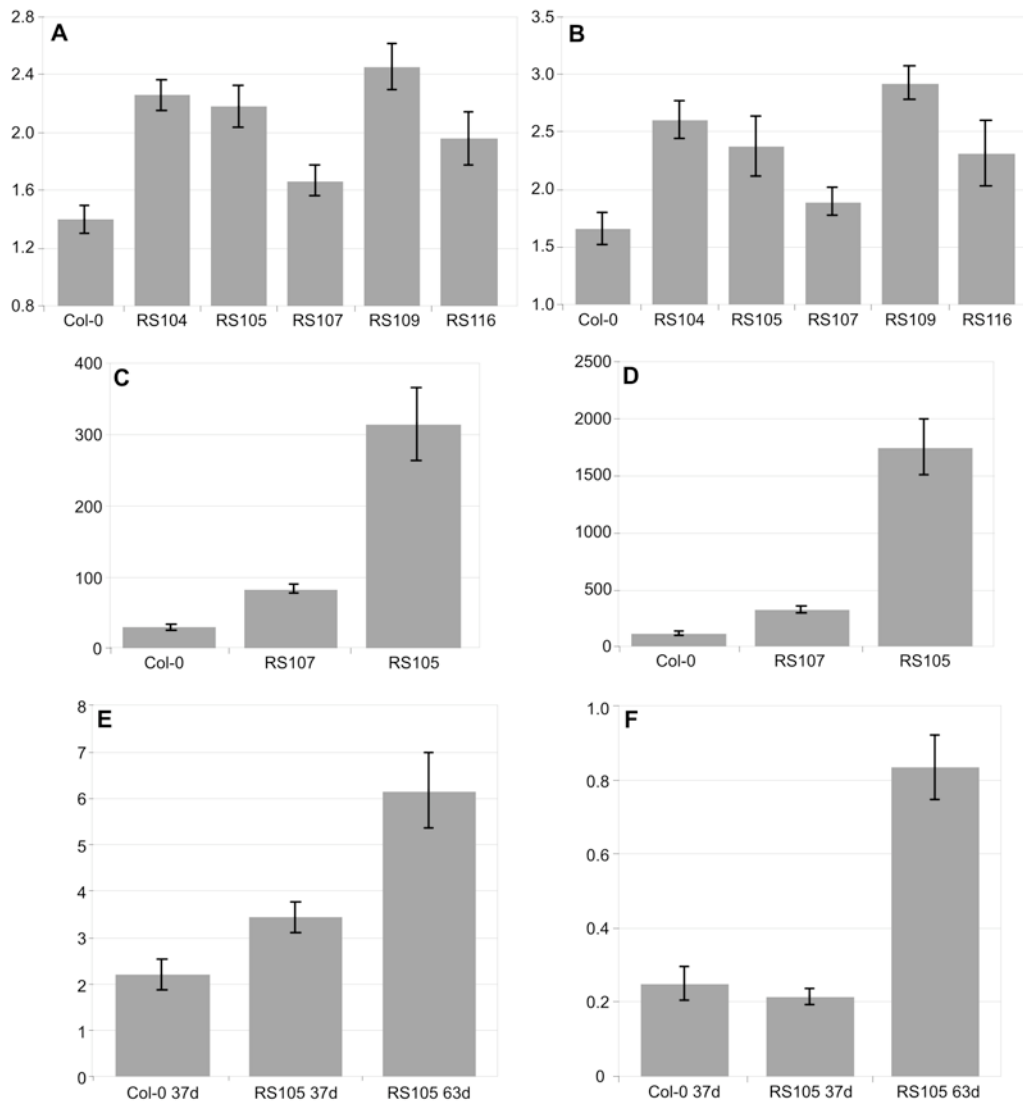
Figure 8 Quantitative analysis of defects in miRNA156 overexpressers

Leaf initiation rates of different miRNA156 overexpresseres in short days **(A)** and long-days **(B)**, averaged from 40 primary transformants. **(C)** Total number of shoots per plant in long-day conditions. **(D)** Total number of leaves per plant in long-day conditions (rosette and cauline leaves). Fresh **(E)** and dry **(F)** weight of miRNA156b overexpressers (in grams) at different stages, averaged from 40 primary transformants.

MiR164 overexpression triggers fusion of lateral organs

CUC1 and *CUC2*, two of the predicted miR164 target genes, redundantly regulate cotyledon separation and maintenance of the shoot apical meristem (SAM), such that double mutants arrest as seedlings and eventually die (Aida et al. 1997). Using tissue

culture, shoots have been regenerated from double mutant calli and produced flowers with fused sepals and stamens (Ishida et al. 2000).

Transgenic plants overexpressing miR164b with predicted increased and ectopic negative regulation of *CUC* genes, phenocopied *cuc1 cuc2* double mutant defects. Strong lines (~1% of primary transformants) displayed fused cotyledons and failed to produce true leaves, while weaker lines (~80% of primary transformants) showed partially fused cotyledons and continued to grow (Figure 9A). Primary shoots fasciated when initiating flower production (Figure 9D), and flowers had fused sepals and stamens, occasionally fewer petals and reduced seed set (Figure 9E). Prolonged attachment of floral organs when senescing was probably due to fused sepals (Figure 9F). In addition to these *CUC* related defects, rosette leaves as well as cauline leaves of miR164b overexpressers were fused along their margins (Figure 9B, C), which was probably related to simultaneous negative regulation of additional *NAC* target genes and mimicked defects observed in *Antirrhinum cup* mutants. These phenotypes were observed with both a long (2539bp) and short (190bp) transgene containing *MIR164b* (see Table 7)*,* and have also been described by other people (Laufs et al. 2004; Mallory et al. 2004a).

Overexpression of *MIR164a*, unlike *MIR164b*, produced only weak vegetative defects, and fusions were mostly limited to floral organs (not shown; (Laufs et al. 2004; Mallory et al. 2004a).

Table 7: MicroRNA164 overexpressing constructs

binary plasmid	MIRNA	size of transgene	sense oligo 5'->3'	antisense oligo 5'->3'
RS110	164a	2225bp	gtggactgaggaggattatac	gggtttaggttttcttcaac
RS111	164b	2539bp	gaccgaaagaatgatggaatg	ctaatagtgatctaaaaggag
RS117	164b	190bp	gaaggtgtgtgatgagcaag	tcaccaaggtggagtggtcatg
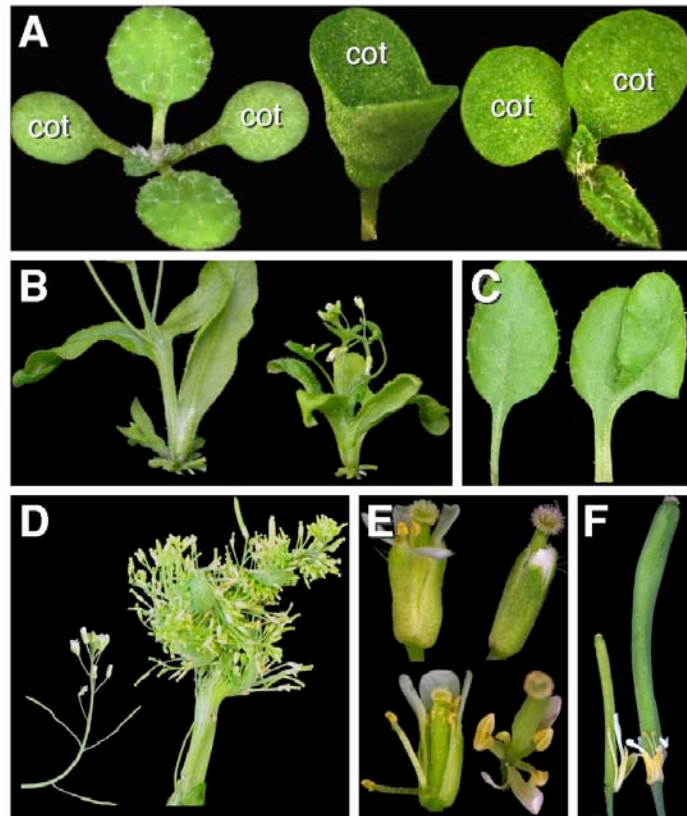
Figure 9: Phenotypes of miRNA164b overexpressers

(**A**) Left, wild-type seedling; middle and right, *35S:miR164b* seedlings with cup-shaped or fused cotyledons (cot). (**B**) Fused rosette and cauline leaves in *35S:miR164b*. (**C**) Left, wild-type leaf; right, two rosette leaves that are fused at the petioles (leaf stems). (**D**) Left, wild-type inflorescence; right, fasciated *35S:miR164b* inflorescence at same magnification. (**E**) Left, wild-type flowers; right, *35S:mir164b* flowers with fused sepals (top) and fused stamens (bottom). At the bottom, some sepals and petals were dissected to reveal stamens. (**F**) Left, latest stage of wild-type development during which senesced sepals, petals and stamens are still attached to the fruit; right, *35S:miR164b* silique, which is much further along than the wild-type siliques on the left, but still has attached senesced organs.

MiRNA172 overexpression results in very early flowering and floral organ transformations

Predicted target genes of miR172 are involved in specification of floral organs (*AP2*; (Bowman et al. 1989)) and regulation of flowering time (*TOE1, TOE2, SMZ, SNZ*; (Aukerman and Sakai 2003; Schmid et al. 2003)). Consistently, overexpression of miR172 in *Arabidopsis* resulted in transgenic plants, which flowered very early with 2-3 rosette leaves in long days and 4-5 in short days (averaged from 40 primary transformants; (Figure 10A,B and 11A, B) with >50% of these lines displaying strong *ap2*-like flower defects (Figure 10C, 11C).

I have overexpressed miR172a, b, and e (see Table 8) and flowering time defects were very similar in all cases. Mir172e, which differs from miR172a and b only at their 5' most nucleotide had lower frequencies of *ap2*-like flowers when overexpressed. This suggests that miR172e is either less effective in negative regulation of target genes or that lower levels of miR172e were produced in overexpressing plants, so that only flowering time, but not flower morphology was affected.

Antisense overexpression of miR172 precursors did not result in any phenotypic changes, as did overexpression of antisense constructs of miRNAs from miR164 or 156 families.
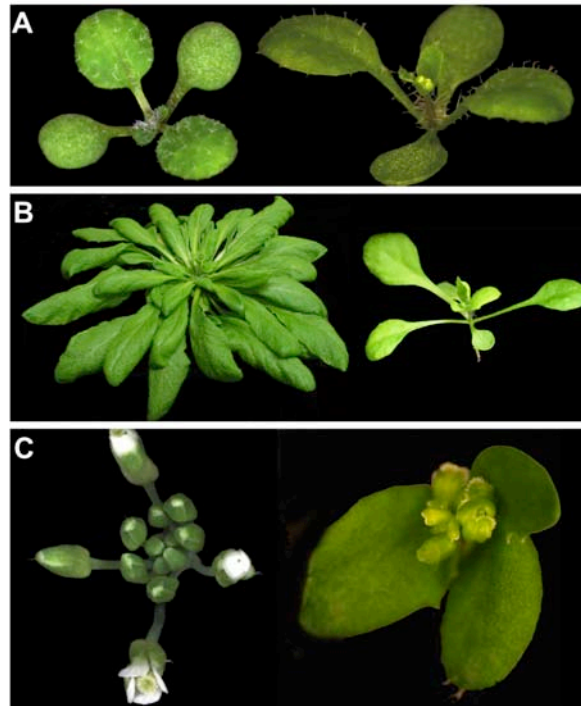
Figure 10: Overexpression of miRNA172 causes early flowering and floral organ transformations

(A) Long-day grown seedlings of similar age. Left Col-0 wild-type, right miR172a overexpresser, which has already started to flower. **(B)** Short-day grown seedlings, which have just started to flower. Left Col-0 wild-type, right miR172b overexpresser. Note the difference in rosette leaf number **(C)** Col-0 (left) and miR172a overexpressing (right) inflorescences.

Table 8 MicroRNA172 overexpressing constructs

binary plasmid	MIRNA	size of transgene	sense oligo 5'->3'	antisense oligo 5'->3'
RS113	172b	2972bp	ctaatgctctcctggtatcgtg	cattcactgctcaaactgtttagg
RS119	172a	411bp	aaaaatggaagactaatttccggag	agcttgtggatctattaatgtcttg
RS120	172b	170bp	tcggcggatccatggaagaaagctc	tttctcaagctttaggtatttgtag
RS133	172a	221bp	aaaaatggaagactaatttccggag	ctgaagaagatctggatggaatcc
RS293	172e	175bp	tgaataggctagcctttggtggatg	gacaagagtagccatgtatttgctg
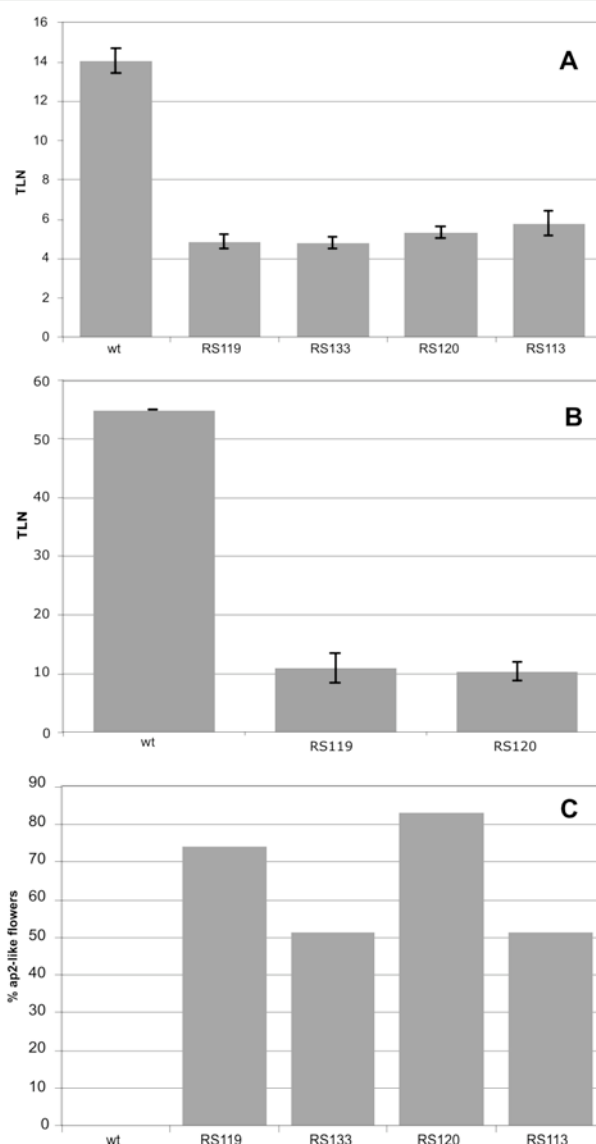
Figure 11: Quantitative analysis of mIRNA172 overexpresser defects

Total leaf number (TLN) of wild-type and different miRNA172 overexpressers in long-days **(A)** and short-days **(B). (C)** Frequency of *ap2-like* floral defects in different miRNA172 overexpressing lines. All numbers are averaged from around 40 primary transformants.

Resistant microRNA targets

MiRNA directed regulation of target genes is mediated by the miRNA complementary target site, which is normally located in the target coding region in plants. Reducing the complementarity to the miRNA by introducing silent mutation in the target site region disrupts miRNA regulation and often results in severe developmental defects (Mallory et al. 2004a; Mallory et al. 2004b; McConnell et al. 2001; Palatnik et al. 2003).

I have mutated miR172 binding sites in *AP2, TOE1* and *TOE3* by overlapping PCR and placed the resulting coding regions behind the CaMV35S promoter. Table 9 lists the resulting constructs.

Table 9: Overexpression of miR172 targets (35S promoter)

binary plasmid	Gene	Identifier	miR172 target site (mismatched nucleotides in bold)
RS190	*wtAP2*	At4g36920	**C**U GCA GCA UCA UCA **G**GA UUC U
RS191	*rAP2*		**C**U GCA GC**C** **AGC** UC**C** **G**GA UUC U
RS196	*wtTOE2*	At5g60120	AU GCA GCA UCA UCA **G**GA UUC U
RS197	*rTOE2*		AU GCA GC**C** **AGC** UC**C** **G**GA UUC U
RS183	*wtTOE3*	At5g67180	**UG** GCA GCA UCA UCA **G**GA UUC U
RS185	*rTOE3*		**UG** GCA GC**C** **AGC** UC**C** **G**GA UUC U
Kpn2I restriction site			TCC GGA
amino acids			A A A S S G F

Primary transformants were grown in long days to determine growth defects.

When overexpressing wild-type forms of *AP2* and *TOE3*, the majority transgenic plants resembled wild-type Col-0, while some *AP2* overexpressers were slightly late flowering or showed *ap2*-like flower defects. While late flowering might be due to high levels of AP2 overcoming miR172 mediated suppression, *ap2* mutant-like phenotypes suggest co-suppression of *AP2* endo- and transgenes (not shown).

Around 10% of primary transformants overexpressing resistant version of *AP2* or *TOE3* (*rAP2, rTOE3*), displayed very similar phenotypes: rosette leaves were dark green and roundish, and the vegetative phase was greatly elongated (Figure 12A). Only about half of the plants ever produced flowers. These were indeterminate in weak lines, such that additional whorls of stamens and carpels were produced

within carpels, which eventually opened to release the extra organs (Figure 12C). Strong lines, did not produce carpels in the center of flowers (Figure 12B), but continuously initiated stamen-resembling organs. None of the transgenic plants produced fertile pollen or seeds.

Resistant *TOE1* overexpressers have not been tested for phenotypic abnormalities.
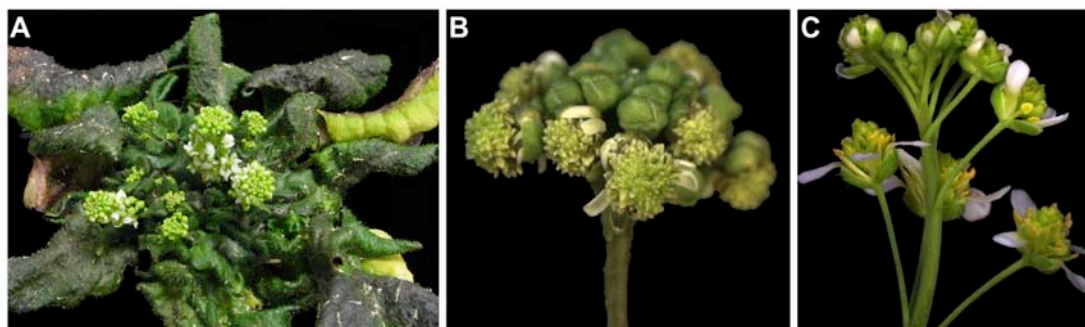
Figure 12: Phenotypes of resistant miRNA target overexpressers

(A) long-day grown plant overexpressing *rAP2* with greatly extended vegetative phase and dark-green rosette leaves. Shoots are initiated simultaneously from several meristems. Inflorescences of *rAP2* **(B)** and *rTOE3* **(C)** overexpressers. Note the additional whorls of stamens in center of flowers.

Conclusions Chapter I

Putative target genes have been predicted for most plant miRNAs based on their small number of mismatches to their miRNAs (Jones-Rhoades and Bartel 2004; Park et al. 2002; Rhoades et al. 2002). I could observe phenotypic defects consistent with targeting of at least some of these genes, for which mutants have been described, by overexpressing miRNAs from three families in *Arabidopsis thaliana.* I also found additional defects in transgenic plants, which were potentially caused by strong and ectopic negative regulation of additional target genes, which have not been functionally characterized, probably due to their redundant nature. Since overexpression of both long and short genomic fragments harboring miRNA stemloops were equally effective in generating developmental defects, sequences outside the stemloop precursor itself are not needed for miRNA production. Most

family members produced very similar defects when overexpressed, suggesting that most known miRNA precursors are functional and produce active miRNAs.

MiRNA156 has been hypothesized to function in flowering time regulation, as its proposed target gene *SPL3* confers early flowering to overexpressing plants when lacking the miRNA binding site in its 3'UTR (Cardon et al. 1997). Overexpressing of miR156 indeed caused late flowering, and I also observed other defects such as reduced apical dominance and increased plastochron, not previously implicated in *SPL* function. Increased side shoot numbers have also been detected when *Arabidopsis MIR156b* (RS105) was overexpressed in tobacco and brassica (Tsegaye Dabi, Salk Institute, unpublished), suggesting that not only miR156 sequence, but also its function has been conserved during evolution. The rice mutant, *plastochron1* (*pla1*) has been identified as a negative regulator of leaf initiation (Itoh et al. 1998). It encodes a cytochrome P450 gene (Miyoshi et al. 2004) and its closest homolog in *Arabidopsis*, *AtCYP78A5*, contains a potential *SPL3* binding site in its promoter region, identical to the one observed in the *AP1* promoter (our unpublished findings). It will be interesting to test, if indirect regulation of *CYP78A5* by miR156 through *SPL* genes mediates the observed plastochron defect in miR156 overexpressers.

Phenotypic analysis of miR164b overexpressers suggests that at least *CUC1* and *CUC2* are indeed regulated by miR164 in vivo, as many phenotypic aspects were overlapping with *cuc1 cuc2* double mutants (Aida et al. 1997). Similar miR164 overexpressing phenotypes have been observed by Mallory et al and Laufs et al (Laufs et al. 2004; Mallory et al. 2004a). Guo et al. (Guo et al. 2005) have focused on miR164 regulation of *NAC1* and they observed decreased numbers of lateral roots in miR164 overexpressers, while both *MIR164a* and *MIR164b* T-DNA insertion mutants showed increased numbers of lateral roots, an effect mediated by *NAC1*. A mutant version of *MIR164c* has been found in a genetic screen by Baker et al (Baker et al. 2005), who observed an increased number of petals in early arising flowers and thus termed the mutant *early extra petals* (*eep*). Since fusions of rosette and cauline leaves, have not been previously described in *Arabidopsis*, it is likely that multiple genes are required to be downregulated simultaneously, unlike in *Antirrhinum*, where a single gene, *CUPULIFORMIS,* regulates separation of organ boundaries (Weir et al. 2004). All reports together suggest a rather broad function of miR164 in *Arabidopsis*, affecting root as well as shoot development.

Early flowering and floral homeotic defects, which I observed in miR172 overexpressers, are consistent with *in vivo* targeting of *AP2* and at least one of the predicted target genes encoding floral repressors. Similar phenotypes have been observed by Chen (Chen 2004), with the exception of miR172e overexpressers, for

which she could not observe changes in floral morphology. Since transformation of Col-0 with my set of overexpresser plasmids yielded stronger defects for miR172a and b when compared to the ones generated by X. Chen (plasmids kindly provided by X. Chen for comparison), I assume that overexpression of miR172e is generally less effective, so that floral defects are less pronounced. Differences in targeting efficiencies due to different 5' terminal nucleotides in miR172a/b and e are not very likely, as many miRNA targets have mismatches to miRNA position 1. Aukerman and Sakai have identified a miR172b overexpressing line termed *EAT-D* from an activation tagging screen for early flowering mutants (Aukerman and Sakai 2003), and this line also displays floral defects resembling *ap2* mutants. Both Chen and Aukerman detected a reduction of AP2 protein abundance in miR172 overexpressers, which was not correlated to reduction of *AP2* transcript levels. Therefore, they postulated that miR172 regulates its targets by translational inhibition, as observed for animal miRNAs, rather than by reduction of transcript levels. Interestingly, mRNA expression of the maize *AP2*-domain gene *GLOSSY15*, which is also regulated my miR172, is inversely correlated to miR172 abundance (Lauter et al. 2005).

Overexpression of miRNAs from the CaMV35S promoter does not only result in increased, but also in ectopic accumulation of small RNAs. The observed phenotypes therefore reflect not only increased miRNA activity in their endogenous domain, but also negative regulation of targets in normally not affected regions. In order to determine, which expression domains are required for observations of distinct phenotypic defects, it will be interesting to drive miRNA expression from different promoters.

A similar problem remains with resistant version of miR172 targets, which I have also overexpressed from the 35S promoter. Expression of miRNA resistant target genes from their endogenous promoter will be required to further define the role or miR172 in spatial regulation of target gene expression. Nevertheless, the overexpressing lines strongly support that miR172 is involved in both floral patterning and flowering time regulation.

Investigation of miRNA expression domains and identification of upstream regulatory proteins has been pioneered by Johnson et al. (Johnson et al. 2003) in *C.elegans* and is under extensive investigation in different model organisms (Johnston and Hobert 2005; Ohler et al. 2004; Parizotto et al. 2004). Analysis of *GUS* fusions for *MIR*156 and *MIR*172 putative regulatory elements resulted in temporal and spatially organized expression domains. Addition of transcribed sequence preceding the miRNA producing hairpin slightly altered the expression in case of

pMIR156c, suggesting that separate elements of regulatory function might be located in this region. For *MIR*172 a, b, and e, putative promoter elements produced very different, yet partially overlapping patterns of expression, which is consistent with the hypothesis of partial redundancy between miRNA family members. The strong expression of *MIR*156 putative promoter elements in stamens and pollen is in line with low expression of *SPL* target genes in these organs. But *MIR172* was not expected to be strongly expressed in petals, since the target AP2 is needed for the specification of these organs. Specification events might however be already completed at the time of miRNA expression in this domain. Nevertheless, the observed patterns need to be confirmed by independent methodologies. Unfortunately, not many techniques are available to distinguish between individual miRNA family members, as most of them rely on hybridization of antisense oligonucleotides to mature miRNAs. However, a recently published report by Chen et al. (Chen et al. 2005) proposes so-called stem-loop RT-PCR to specifically PCR amplify individual precursors from cDNA. In order to draw conclusive statements on redundancy of *MIRNA* family members from these expression studies, me or others will have to extend this analysis to additional family members and also analyze individual members in more detail, as for example by extending the presumptive regulatory element in the 3' direction, which has already slightly altered *GUS* patterns in case of miR156c.

An interesting application of stable promoter-*GUS* lines will be the identification of putative upstream regulators by crosses of candidate regulators to both reporter and overexpressing lines. Similarly, reporter lines can be mutagenized and screened for mutants with altered pattern of reporter gene expression.

CHAPTER II

DETERMINANTS OF microRNA:TARGET RECOGNITION

This work was published in Developmental Cell (2005) Vol.8, 517-527.

Specific Effects of MicroRNAs on the Plant Transcriptome

Rebecca Schwab*, Javier F. Palatnik*, Markus Riester, Carla Schommer, Markus Schmid and Detlef Weigel

* These authors contributed equally

Contributions to this chapter:

Most of the experiments described in this chapter have been carried out by myself, often in collaboration with JP, and some were done by other lab members without my help. Since all parts are necessary to understand the main findings described here, I included them all and specifically indicated experiments, in which I was not involved. In addition, I describe all findings from the perspective of the group of people involved. The text is partially adapted from the manuscript.

JP and CS generated and phenotyped miRNA159 and 319 overexpressers and hybridized their RNA to microarrays
RS did microarrays for miR156, 164, and 172 overexpressers with help by MS
RS analyzed microarrays with help by JP and DW, MR and DW carried out statistical tests.
RS developed miRNA targeting rules based on microarray data, as proposed by JP. A computational tool to assist this analysis was developed by MR
RS carried out RACE-PCRs to determine cleavage products of miRNA targets
JP designed the quantitative RACE-PCR experiment, which was carried out by RS
RS analyzed miR172 feedback regulation, using methods trained by JP
DW was leading the course of all experiments and discussed results.
RS and DW wrote the manuscript, which was used as a basis for this chapter.

Plant miRNAs are thought to function on the RNA level, since cleavage products consistent with small RNA mediated processing have been isolated from multiple target genes (Kasschau et al. 2003; Llave et al. 2002b). Direct evidence for a plant miRNA functioning by reducing target transcript levels has first been generated by expression profiling of miRNA319a overproducing plants using quantitative RT-PCR and microarrays (Palatnik et al. 2003). Consistent with the idea of cleavage mediated transcript degradation, exonucleases have been implicated in degradation of cleavage products (Souret et al. 2004). MiRNA172, however, has been suggested to rather function by translational inhibition, the general mode of animal miRNA action, as changes in target levels were only apparent on the protein level (Aukerman and Sakai 2003; Chen 2004).

In order to test if other miRNAs also perform their repressive activity by transcript degradation, we studied their effects on target transcript levels in miRNA overexpresseres by RT-PCR and Affymetrix ATH1 microarrays. In this analysis, we included not only miRNAs 156b (transgene RS105) and 164b (transgene RS117), but also miRNAs 159a and 319a. (Overexpressers of miRNA 159a have been generated, phenotyped and hybridized to Affymetrix arrays by Javier Palatnik, and the activation tagged miR319a overproducer *jaw-D* has been processed by Carla Schommer.) MiR159 has been predicted to regulate 8 GAMYB transcription factors (Table 10) (Rhoades et al. 2002), which are expressed rather ubiquitously (MYB33 and MYB65), or preferentially in stamens (MYB101, MYB120) (Figure 13). Consistently, overexpression of miR159 from the 35S promoter led to stamen defects and male sterility. Delayed flowering has been reported as an additional phenotype in a similar experiment carried out by Achard and colleagues (Achard et al. 2004). MiR319a, also known as miR-*JAW*, regulates five *TCP* transcription factors (see Table 10) which are involved in leaf morphogenesis (Palatnik et al. 2003). Overexpressers of miR319a show epinastic cotyledons, crincly leaves and slightly delayed floral transition. Microarrays of vegetative apices from miR319a overexpressers have been analyzed previously (Palatnik et al. 2003) and we added leaves as a separate tissue with high target gene expression.

We made use of the AtGenExpress expression atlas, which is also based on Affymetrix ATH1 microarrays (Schmid et al. 2003) (www.weigelworld.org/ resources/microarray/AtGenExpress/) to determine tissues with maximal expression of predicted target genes. Figure 13 shows expression profiles of target genes in different tissues during wild-type development. For miRNAs 164b and 156b we focused on young inflorescence tissue. As some of the predicted miR164 targets were absent in apex tissue, but preferentially expressed in roots, we determined their

expression changes by quantitative RT-PCR using root tissue of wild-type and miR164b overexpressing plants. Open flowers were used in case of miR159.
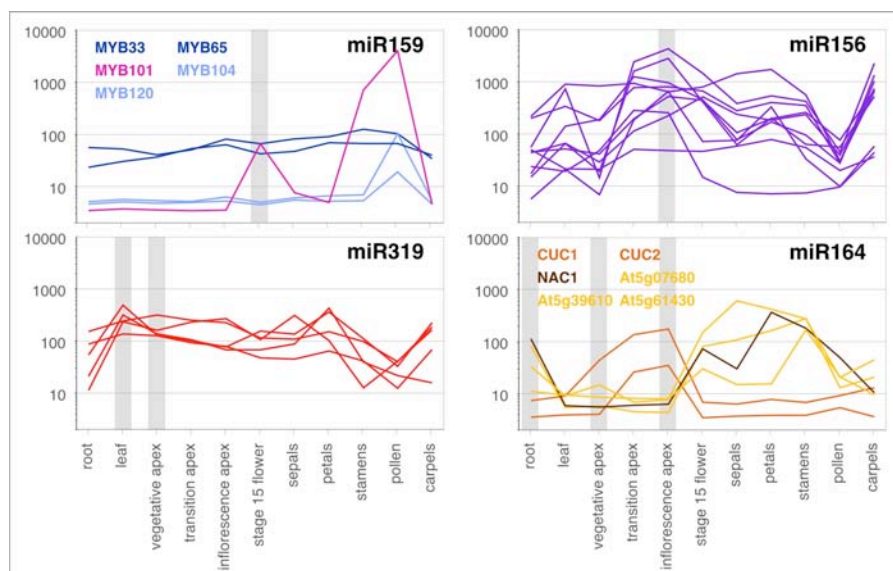
Figure 13: Expression patterns of predicted miRNA target genes in wild-type

The miRNA whose targets are presented is always shown on the top right. Expression estimates by gcRMA are from the AtGenExpress expression atlas. Gray bars indicate tissues analyzed in miRNA overexpressers. Floral organs are from stage 15 flowers.

Table 10: MicroRNA binding elements in predicted target genes of microRNAs 159 and 319

Gene name	Identifier	Target sequence 5'->3' (mismatches to miR159a/319a in bold)	Number of mismatches
MYB33	At5g06100	U**GG** AGC UCC CUU CA**U** UCC AA**U**	3
MYB65	At3g11440	U**G**G AGC UCC CUU CA**U** UCC AA**U**	3
MYB81	At2g26960	U**CG** AG**U** UCC CUU CA**U** UCC AA**U**	4
MYB97	At4g26930	**AU**G AGC UC**U** CUU CAA **A**CC AAA	4
MYB101	Az2g32460	UAG AGC U**UC** CUU CAA **A**CC AAA	2
MYB104	At2g26950	U**GG** AGC UCC CUU CA**U** UCC AA**G**	3
MYB120	At5g55020	**AGC** AGC UCC CUU CAA **A**CC AAA	4
MYB125	At3g60460	U**GG** AGC UCC **A**UU C**G**A UCC AAA	3
TCP2	At4g18390	**A**GG **GGG A**CC CUU CAG UCC AA	4
TCP3	At1g53230	**A**GG **GGU C**CC CUU CAG UCC A**U**	5
TCP4	At3g15030	**A**GG **GGU C**CC CUU CAG UCC A**G**	5
TCP10	At2g31070	**A**GG **GGU A**CC CUU CAG UCC A**G**	5
TCP24	At1g30210	**A**GG **GGG A**CC CUU CAG UCC AA	4

Effects of microRNA overexpression on predicted target genes

Specific effects of miRNA overexpression were particularly apparent in miR156 overexpressers: of the 15 *SPL* genes represented on the affymetrix ATH1 array, 10 genes with predicted miR156 target sites were substantially reduced in expression, while the remaining genes without miR156 target sites were unaffected (Figure 14).

Six *NAC* family members have been predicted as targets of miR164, including *At5g39610* with 4 mismatches (Jones-Rhoades and Bartel 2004). It has been suggested that only four of these are efficiently guided to cleavage by miR164 (Laufs et al. 2004), although cleavage site mapping in wild type is consistent with all of them being miR164 targets (Figure 15; (Mallory et al. 2004a)). *CUC1* and *CUC2* were significantly reduced in microarrays on inflorescence tissue. Using quantitative RT-PCR, we found that expression of the remaining four genes, including *At5g39610*, was substantially reduced in roots of miR164 overexpressers (Figure 14).

The miR159 target *MYB33* has previously been shown to be downregulated in leaves of miR159 overexpressers (Achard et al. 2004). Surprisingly, we did not detect a change in *MYB33* or *MYB65* RNA expression in open flowers, even though both genes were easily detected in the controls. The arrays showed only one predicted target, *MYB101*, to be significantly reduced. Quantitative RT-PCR carried out by Javier Palatnik revealed that another target, *MYB120*, which was not detected on the arrays, was downregulated as well (not shown) While these observations confirm that miR159a is capable of guiding target RNA cleavage, they also suggest that higher levels of miR159a than present in our overexpressing lines are required for significant reduction of *MYB33* and *MYB65* mRNAs. Notably, *MYB33* and *MYB65* transcript abundance is much higher and more even throughout development than that of the other miR159 targets (Figure 13), suggesting that miR159 and its targets *MYB33* and *MYB65* are closer to equilibrium than other pairs of miRNAs and targets.

Similar to previous experiments with vegetative apices (Palatnik et al. 2003), all five *TCP* target genes were substantially reduced in leaves when miR319a was overexpressed (Figure 14).

Taken together, this analysis supports that miRNAs 156, 159, 164 and 319 indeed regulate their predicted targets – *SPL, NAC, GAMYB*, and *TCP* transcription factor genes respectively – by cleavage induced reduction of RNA levels.
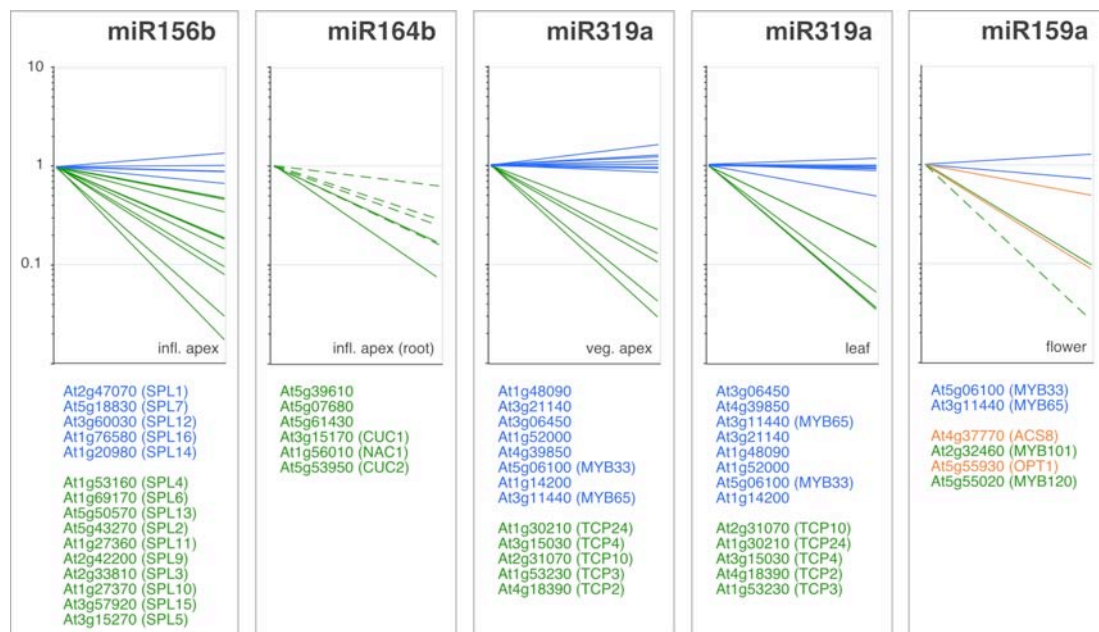
Figure 14: Expression analysis of miRNA overexpressing plants

Gene identifiers are in the same order from top to bottom as the lines in the graphs above. Wild type is always on the left, and miRNA overexpressers on the right; expression values are normalized to the wild-type control. Solid lines indicate microarray data, dashed lines real-time RT-PCR data. Genes that change significantly in expression are indicated in green. Blue indicates genes that do not change significantly and include the following: for miR156b, *SPL* genes that do not contain miR156 target sites; for miR319a, 8 genes with 4 mismatches; for miR159a, two *MYB* genes related to the genes that are significantly downregulated. For miR159a overexpressers, ochre indicates two genes with 4 mismatches that are not conserved targets, but are significantly downregulated. MiR164b data are from inflorescence apices (*CUC1* and *CUC2*), the remainder from roots.

Summary of published computational prediction of miRNA target genes

There have been several attempts to identify plant miRNA targets from first principles. Rhoades and colleagues (2002) initially predicted targets based on the observation that there were statistically significantly more mRNAs with 3 or fewer mismatches to authentic miRNAs than to randomized miRNAs. These predictions have been refined by, for example, counting G:U pairs as 0.5 mismatches. Additional genes are also considered as targets if they belong to the same gene family as a previously predicted target and share sequences closely related to the miRNA complementary motif (Jones-Rhoades and Bartel 2004). Even so, the *TCP* genes only conform to these predictions if one treats the closely related miRNAs miR159 and miR319 as interchangeable, which is inconsistent with in vivo data (Achard et al. 2004; Palatnik et al. 2003) (Palatnik et al., submitted). At the same time, genes

sharing the same mismatch limitations are generally not considered as targets, if the presumptive target sites are not conserved in other species (Jones-Rhoades and Bartel 2004).

Since microarrays have successfully confirmed predicted targets of miRNAs 156, 159, 164 and 319, we wanted to extend our studies and systematically analyze effects of miRNA overexpression on the remainder of the genome in order to determine if predicted target genes represent the full extend of direct targets for these miRNAs

Effects of miRNA overexpression on genes with high sequence complementarity

As our microarray analysis was carried out on plants constitutively overexpressing miRNAs, a large fraction of the transcriptome was downregulated when compared to wild-type controls, especially because direct targets function as transcriptional regulators. In order to separate primary from secondary effects, we initially limited potential direct targets to those with considerable sequence complementarity to the overexpressed miRNA.

We first determined all transcripts with up to 4 mismatches to any of the 4 miRNAs, and found that these were enriched among downregulated genes (Table 11; statistical significance cannot be determined because of the small numbers of expected and observed cases). Excluding previously predicted targets from the group of analyzed genes showed that the excess of observed downregulated genes was readily explained by the presence of known or previously predicted targets and no more genes than expected by chance were downregulated among additional genes with up to 4 mismatches. Only in the case of miR159, two additional downregulated genes were found in the class of genes with up to 4 mismatches (Table 11).

These were At5g55930, which encodes an oligopeptide transporter (OPT1) (Koh et al. 2002), and At4g37770, a 1-aminocyclopropane-1-carboxylate synthase (ACS8). Both genes were unrelated to the previously identified *MYB* targets. We decided to focus on *OPT1*, because the reduction in RNA expression was particularly dramatic and very similar to the reduction seen for *MYB101* (Figure 14). Using a previously established RACE-PCR protocol for detecting miRNA-guided cleavage products (Kasschau et al. 2003) (see Materials and Methods), we found that in flowers of miR159a overexpressers, *OPT1* mRNA was cleaved at a site opposite of

58

position 10 to 11 of the miRNA (Figure 15), which is typical for miRNA- or siRNA-guided cleavage (Elbashir et al. 2001b; Kasschau et al. 2003). We did not detect any cleavage products in wild type, where *OPT1* expression is highest expressed in pollen, similar to *MYB101* expression (Figure 16). These findings suggest that the *OPT1* transcription pattern in flowers does not normally overlap with miR159 expression. Alternatively, higher miR159 levels than present in wild type are required for *OPT1* cleavage.
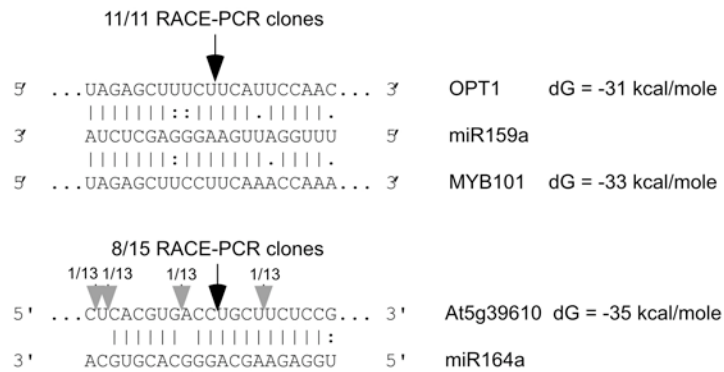

Figure 15: Cleavage site mapping

The 5' ends of *OPT1* and At5g39610 cleavage products, mapped by RACE-PCR, are indicated by an arrow.
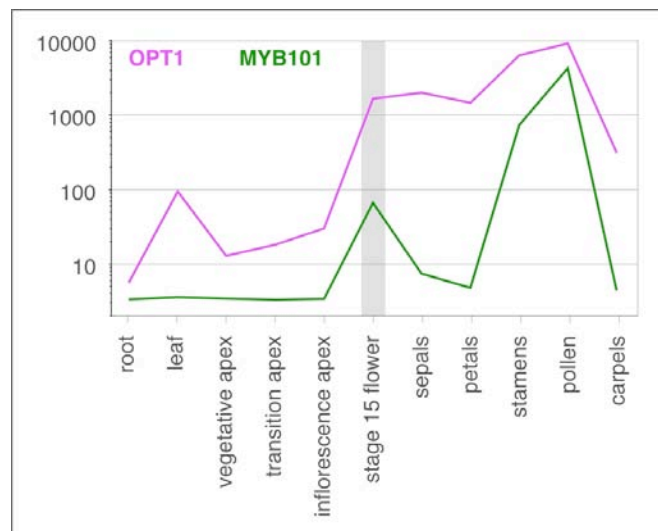
Figure 16: Expression of *OPT1* in wild type

Expression estimates by gcRMA are from the AtGenExpress expression atlas.

Table 11: Summary of downregulated genes in four miRNA overexpressers

		≤3	Mismatches 4	5	Perfect match pos. 2-12
Present in control		17	31	299	87
Observed downregulated	all	13	4	7	18
	conserved targets	13	2	3	15
	others	0	2	4	3
Expected downregulated without conserved targets		0	1	8	3

Across all four comparisons, 59,395 genes were called present in the controls, of which 1,906 (3.2%) were significantly downregulated in the miRNA overexpressers. For miR164b and miR319a, only one tissue was included in this summary (inflorescence and vegetative apices, respectively).

Effects of miRNA overexpression on genes with limited sequence complementarity

Since *OPT1* has not previously been predicted as a miR159 target and *TCP* genes would not have been predicted by previous algorithms as miR319 targets, we next asked if additional genes with greater number of mismatches were also directly regulated by our set of miRNAs.

It has been reported for siRNAs, which act in a similar manner as most plant miRNAs, that perfect complementarity to siRNA positions 2 to 12 can be sufficient to trigger cleavage in vitro (Haley and Zamore 2004). Based on this finding, it has been predicted that artificial siRNAs have substantial effects not only on their intended targets, to which they are complementary in their entire sequence, but also on other mRNAs with more limited complementarity. Downregulation of such so-called off-targets has been confirmed by expression profiling of cultured human cells injected with synthetic siRNAs (Jackson et al. 2003)

To detect effects of plant miRNAs on genes with limited sequence complementarity, we first analyzed genes that were perfectly complementary to nucleotides 2 to 12 of the miRNAs, but were not conserved targets. Across the four overexpressers, 3.2% of genes that met these conditions were downregulated in response to the corresponding miRNA, which was the same as expected by chance (Table 11).

Because the first eight nucleotides of the miRNA have been considered as the core region of the miRNA-target sequence interaction in animals (Doench and

Sharp 2004), we next focused on potential targets that satisfied this criterion. Among genes with 5 mismatches to the respective miRNAs, there were 16 genes with perfect matches to positions 1 to 8 of the miRNAs under investigation. Of these, only one gene was affected by miRNA overexpression, At4g37770 (ACS), which we had already identified as a putative miR159a target (see above).

Analyzing the effects of a common hexamer in the 5' part of the miRNAs (seed region), which is sufficient in animals for target regulation (Brennecke et al. 2005, Lim et al. 2005), we determined all transcripts with perfect matches to position 2-7 or 3-8 to the four miRNAs analyzed. As seen before, only minor differences between observed and expected downregulated genes have been found, and these were not statistically significant (not shown).

We also examined all genes with up to five mismatches that were downregulated in each overexpresser. Of a total of 24 genes with 5 or fewer mismatches that were downregulated in four overexpressers, only 6 genes were not conserved targets (expected: 9; no significant difference under chi-square test; Table 11), so that not more genes than expexted by change were downregulated. These general findings did not change when we counted G:U only as 0.5 mismatch (not shown).

Finally, we asked whether downregulated genes that were not conserved targets were overall more similar in sequence to the overexpressed miRNA than random genes (analysis carried out by Markus Riester). As a measure of sequence similarity, we used Smith-Waterman (1981) scores, and compared downregulated genes with all genes present in the control. For all four miRNAs, the downregulated genes were on average less, rather than more similar to the miRNA, although this difference was not statistically significant (Figure 17; Table 12).

In conclusion, by several different measures, we could not find evidence for major transcriptional off-target effects of plant miRNAs. Apart from the conserved targets, the only other mRNAs affected directly by miRNA overexpression seem to come from genes such as *OPT1*, whose sequence complementarity in the miRNA target site is as high as that of conserved targets. This suggests that the extend of plant miRNA targets is rather small and normally limited to a small set of related genes, which are often evolutionary conserved. Additionally this finding makes it likely that not only the evolution of miRNAs is an ongoing process (Allen et al. 2004), but also that of miRNA targets
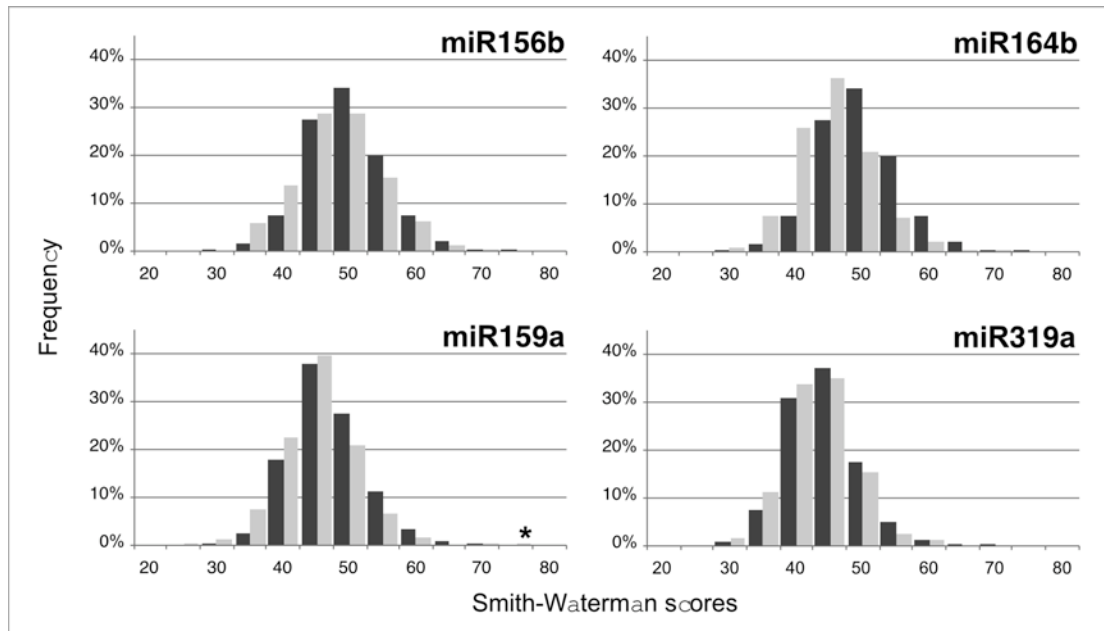
Figure 17: Absence of major off-target effects of plant miRNAs (by Markus Riester)

Distributions of Smith-Waterman (1981) scores are similar between genes that are significantly downregulated in response to miRNA overexpression (light gray), and all genes present in the control (dark gray), once authentic, conserved targets are excluded. For miR159a, asterisk indicates At5g55930 (*OPT1*), which is guided to cleavage by miR159a (Figure 15).

Table 12: Means of Smith-Waterman scores (Determined my Markus Riester)

Experiment	Tissue	Control mean	Experimental mean	t	p-value
35S::miR156b	Infl. apex	47.66455	45.76623	2.397474	0.991457
35S::miR164b	Infl. apex	47.66455	42.83555	16.07713	1.000000
35S::miR159a	Flowers	44.83523	42.88953	2.965378	0.998387
35S::miR319a	Infl. apex	42.16652	41.08654	0.311017	0.622032
35S::miR319a	Leaves	42.19106	41.41667	-0.425998	0.335529

All genes present in control were compared with genes significantly downregulated in response to miRNA overexpression. Student's t-tests indicate that the means of the experimental samples are not significantly different than those of the controls (the actual experimental means are lower than the control means).

62

Sequence requirements for miRNA:target recognition

The previous finding that miR319a targets several *TCP* genes with 4 or 5 mismatches was the first indication that the initial boundary condition of maximally 3 mismatches proposed by Rhoades and colleagues (Rhoades et al. 2002) was too narrow (Palatnik et al. 2003). When counting G:U pairs only as 0.5 mismatch, *TCP3* still has 4 mismatches, and exceeds the 3.5 mismatch limit proposed by Jones-Rhoades and Bartel (Jones-Rhoades and Bartel 2004) for miRNA target identification (see Table 10)

To identify parameters that unambiguously distinguish targets from non-targets regardless of sequence conservation in other species or gene family members, we analyzed genes with up to 5 mismatches to each of the four miRNAs. There were 317 potential target sequences in 305 genes that were detected in the wild-type controls (176 for miR156, 47 for miR159, 29 for miR164, and 65 for miR319). Among these sets, we compared those genes that were not affected by miRNA overexpression ("non-functional sequences") with those that were ("functional sequences").

Free energy of miRNAs paired with potential target sites. We aligned potential target sequences with the corresponding miRNA and calculated their free energy (dG) (Zuker 2003). When we sorted the potential target sites according to free energy, the functional sequences always ranked among the top of the lists, but free energy alone did not distinguish functional and non-functional sequences (Table 13).

An analysis of animal miRNA targets sites has shown that interaction of target RNAs with the first eight nucleotides of the miRNA is crucial for translational inhibition in an in vitro system (Doench and Sharp 2004). We therefore sorted potential target sequences also according to their free energy when paired with the first eight nucleotides of our set of plant miRNAs. This resulted in validated targets appearing at the top of the respective lists again, but as with the complete miRNAs, this procedure did not unambiguously discriminate between targets and non-targets.

Table 13: Ranking of top seven genes by free energy

MicroRNA	Gene	Free energy	
		kcal/mole	% of maximum
miR156b	At2g33810 (SPL3)	-36.6	88.8
	At5g50570 (SPL13)	-35.7	86.7
	At5g43270 (SPL2)	-35.7	86.7
	At3g57910 (SPL15)	-35.7	86.7
	At2g42200 (SPL9)	-35.7	86.7
	At1g69170 (SPL6)	-35.7	86.7
	At1g27370 (SPL10)	-35.7	86.7
miR159a	At2g26950 (MYB104)	-38.1	93.8
	At5g06100 (MYB33)	-37.8	93.1
	At3g11440 (MYB65)	-37.8	93.1
	At3g60460	-34.2	84.2
	At2g26960 (MYB81)	-32.8	80.8
	At2g32460 (MYB101)	-32.6	80.3
	At3g06450	-32.3	79.6
miR164b	At5g53950 (CUC2)	-42.7	88.0
	At3g15170 (CUC1)	-42.7	88.0
	At5g07680	-41.9	86.4
	At1g56010 (NAC1)	-41.8	86.2
	At5g61430	-41.7	86.0
	At1g57820	-37.9	78.1
	At2g47650	-37.7	77.7
miR319a	At5g06100 (MYB33)	-37.3	83.1
	At3g11440 (MYB65)	-37.3	83.1
	At2g26950 (MYB104)	-37.3	83.1
	At1g31880	-35.3	78.6
	At3g15030 (TCP4)	-35.0	78.0
	At2g31070 (TCP10)	-35.0	78.0
	At4g18390 (TCP2)	-35.0	78.0

Targets are indicated in green, non-targets in black. Free energies were calculated with mfold (Zuker 2003)

Position of mismatches. The free energy of miRNA:target interaction is influenced by the relative position of mismatches, both within the duplex and with respect to each other. Most functional miRNA target sequences had long stretches of perfectly matching nucleotides, especially to the 5' portion of the miRNA. In contrast, runs of internal mismatches were limited to two contiguous mismatches (Figure 18A, B).

64

We found several non-targets that had free energies similar to functional targets, but had mismatches at positions 10 or 11, which flank the cleavage site in functional miRNA targets (Figure 18C). We conclude that mismatches at these positions are not allowed. This contrasts with recent reports that suggested allowable mismatches at position 11 (Allen et al. 2004; Mallory et al. 2004b). Reports on siRNAs in animals are in line with these positions being particularly important, even though there are some cases in which individual mismatches are tolerated (Boutla et al. 2001; Elbashir et al. 2001b; Holen et al. 2002).

Consistent with experimental analysis of animal targets and mutational analysis of plant targets (Doench and Sharp 2004; Laufs et al. 2004; Lewis et al. 2003; Mallory et al. 2004b; Parizotto et al. 2004; Vaucheret et al. 2004), we found the region pairing with the 5' portion of the miRNA to be specifically sensitive to mismatches. We can extend the mismatch-sensitive region for our examples of plant miRNAs from positions 2 to 8, which are critical for animal miRNAs, to positions 2 to 12, as we never found a functional target sequence with more than one mismatch in this region. Even *TCP* genes with a total of 5 mismatches are perfectly complementary to miR319a in this region (Figure 18A). Several genes that had otherwise low free energy, including no mismatches at the presumptive cleavage site and no runs of more than two contiguous mismatches in the region complementary to the 3' end of the miRNA, were insensitive to miRNA overexpression, if they had mismatches in the 5' region (Figure 18D). Conversely, even if a potential target sequence had 10 or more consecutive matches in the 5' region, it was not functional, if there was a stretch of three mismatches toward the 3' portion of the miRNA (Figure 18B).

There remain a few genes that did not behave as inferred from the rules discussed above. This set comprises *MYB33* and *MYB65*, which appear to be less efficiently cleaved than *MYB101* by miR159a, even though *MYB101* has a higher free energy. It seems unlikely that sequences outside the target site are important in this context, since miR159a efficiently targets *OPT1*, which is unrelated to *MYB101* outside of the target sequence. Similarly, *MYB33* and *MYB65* are unaffected by miR319a, which is related to miR159a, but targets *TCP* genes. The major difference between these groups is located in the region pairing with the 5' half of the miRNA. *MYB101* and its paralog *MYB120* have mismatches to position 6 of miR159a, while *MYB33* and *MYB65* have mismatches to position 7, and one to position 1. We can therefore not exclude that the specific position of single mismatches in the 5' region is of particular importance, although there are alternative explanations for the
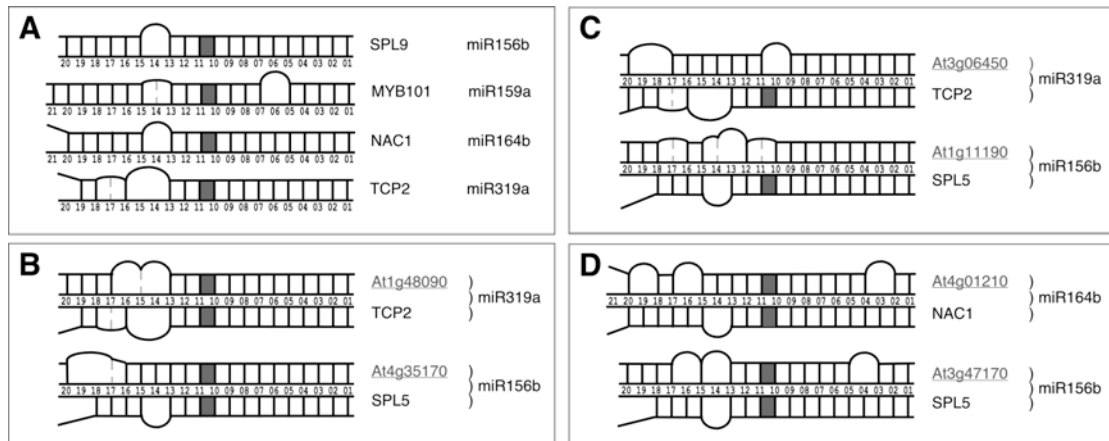
Figure 18: Features that determine miRNA target selection

(**A**) Validated targets; numbers refer to miRNA nucleotide positions, the cleavage site between position 10 and 11 is marked in gray. Dashed gray lines indicate G:U mismatches. (**B-D**) Comparison of miRNA interaction with non-targets (top, underlined, gray) and authentic targets (bottom). (**B**) Requirement of internal stretches of less than three contiguous mismatches, even when region complementary to 5' end of miRNA matches perfectly. (**C**) Requirement of perfect matches surrounding potential cleavage site. (**D**) Mismatches in region complementary to 5' end of miRNA often have adverse effects on targeting.

insensitivity of *MYB33* and *MYB65* to miR159 overexpression in our experiments, as discussed above.

Summary of empirical rules for target recognition. Authentic miRNA targets sites are distinguished by low overall free energy when paired with the corresponding miRNA (in the set examined, at least 72% compared to perfect match). Only one mismatch is tolerated in the region complementary to nucleotides 2 to 12 of the miRNA, but not at the cleavage site. Information outside the miRNA target sequence does not seem to be required for efficient transcript cleavage, as indicated by the *MYB* genes and the unrelated *OPT1* gene, both of which are guided to cleavage by miR159a. A similar case is provided by the two classes of miR156 targets. Although both comprise *SPL* genes, one group has the target motif in the coding region, while the other has the target motif in the 3' UTR, and there is no extended sequence similarity outside the miRNA target site.

A final issue that should be taken into account is the relative ratio of a miRNA and its target (Doench and Sharp 2004). As discussed above, the higher basal levels of *MYB33* and *MYB65* expression may explain why these genes respond very little to miR159a, whereas *MYB101* and *MYB120* respond strongly, even though the latter genes have higher free energy than *MYB33* and *MYB65* when paired with miR159a.

66

Validation of miRNA target rules with other known targets

In a first step, we randomized the four miRNA sequences, an approach that has been used to estimate the power of target predictions based on sequence complementarity (Jones-Rhoades and Bartel 2004; Rhoades et al. 2002). In comparison to previously used rules (Jones-Rhoades and Bartel 2004), the experimentally inferred parameters produce fewer hits with randomized miRNA sequences (done by Markus Riester, not shown).

Our next step was to determine whether these parameters also applied to other miRNA-target pairs. We focused initially on miR160, miR165/166, miR167, miR168, miR170/171, miR393, miR394, miR395 and miR397, which can be aligned with their targets without bulges. We confirmed that these miRNA:target pairs have low free energies, no mismatches at the cleavage sites, long stretches of perfect matches to the 5' portion of the miRNA, and no strings of more than two mismatches to the 3' portion of the miRNA. *AGO1*, which is targeted by miR168 (Vaucheret et al. 2004), is the only target with two mismatches between positions 2 and 12, and only one additional mismatch in the miRNA 3' part. The same rules also apply to targets of miR162 and miR396, if one considers a bulge in the mRNA as a mismatch, as well as two of the three proposed miR398 targets. (The bulges, if they are present, are always found opposite of the 5' region of the miRNA). Only the miR398 target *CSD2*, with two G:U wobbles and a bulge in nucleotides 2 to 12, is an exception (Jones-Rhoades and Bartel 2004). However, it needs to be confirmed that these predicted targets are indeed regulated by miR398, and not a closely related miRNA. In addition, it is unknown how efficient the miRNA-induced cleavage is. The small number of exceptions demonstrates that the rules that we deduced from the analysis of four miRNA overexpressers are broadly applicable.

As plant miRNAs are similar to siRNAs in their mode of action, we compared highly potent siRNAs, as defined by Reynolds et al. (Reynolds et al. 2004), with all plant miRNAs known or proposed to cause cleavage. Most of the criteria for siRNAs applied also to plant miRNAs, among them low internal stability at the 5' end of the small RNA, which can be partially attributed to U being the most common nucleotide at position 1. Whereas the average overall GC content of the analyzed 19 miRNA families is 53.4%, it is 42.9% for the first 5 nucleotides, with only the 5' end of miR394 having a higher stability than the entire miRNA. Among specific sequence biases, Reynolds and colleagues (Reynolds et al. 2004) found an A at position 10 to be the most important one. 50.9% of plant miRNAs (adjusted for frequency within a

family) have an A at this position. Since there is never a mismatch, the target always has a U at this position, consistent with endonucleases preferring to cleave 3' of a U (Donis-Keller 1979; Reynolds et al. 2004).

Analysis of miR172 targets

While cleavage seems to be the predominant mode of plant miRNA action, miR172 has been reported to function primarily by translational inhibition (Aukerman and Sakai 2003; Chen 2004). Predicted targets of miR172 are *APETALA2* (*AP2*) along with five other members of the *AP2* family, *TOE1*, *TOE2*, *TOE3*, *SMZ* and *SNZ* (Park et al. 2002; Schmid et al. 2003). Our rules for miRNA targets would predict that these genes are subject to miR172-guided cleavage, since the predicted targets have only one mismatch in the region complementary to nucleotides 2 to 12 of the miRNA, and at most three other mismatches outside this region (ses Table 3). Although products consistent with miR172-guided cleavage can be detected in wild-type plants for all of these genes (Aukerman and Sakai 2003; Kasschau et al. 2003) (J. Mathieu, unpublished), it has been suggested that these cleavage products are very rare and that they do not contribute to normal miR172 function (Aukerman and Sakai 2003; Chen 2004).
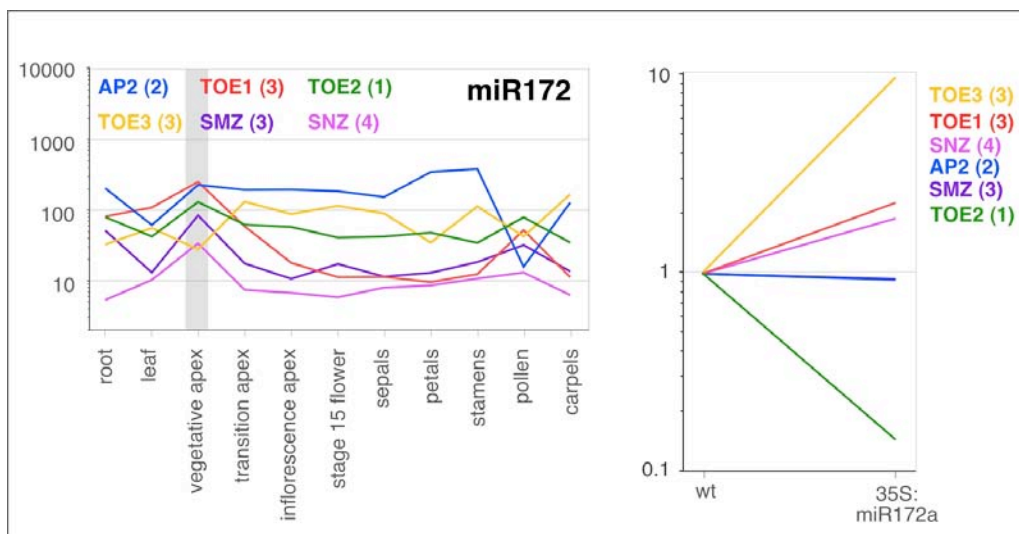
**Figure 19: Expresion of miR172 targets in wild-type and miR172 overexpressers**

Left: Expression of miR172 target genes in different tissues during wild-type development. Expression estimated by gcRMA are from the AtGenExpress expression atlas. Right: Response of predicted miR172 targets to miR172a overexpression in vegetative apices. The apparent increase in *SNZ* expression was not statistically significant.

We used Affymetrix arrays to analyze vegetative apices of miR172 overexpressing plants, with highest abundance of target gene expression in wild-type (Figure 19). Only a single gene with up to 5 mismatches to miR172a was significantly downregulated in miR172a overexpressing plants: *TOE2* (Figure 19).

Because the arrays measure only steady-state levels of mRNA, we wanted to confirm that the substantially reduced levels of *TOE2* were indeed due to increased cleavage. To this end, we (Javier Palatnik) developed an RT-PCR-based assay for quantification of miRNA-derived cleavage products (Figure 20). Therefore, we generated cDNA libraries with adapters ligated to the 5' end of all non-capped transcripts (which should include cleavage products, see Materials and Methods) (Kasschau et al. 2003). Specific miRNA cleavage products were amplified with forward oligonucleotides spanning the adapter-transcript boundary and gene-specific reverse oligonucleotides.
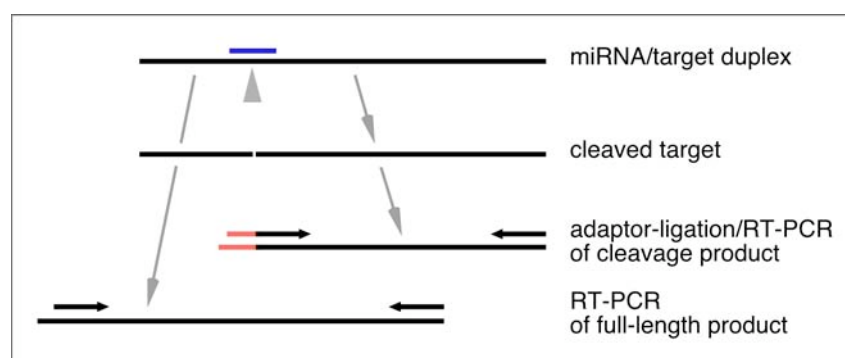
Figure 20: PCR assay for quantification of cleavage products

MiRNA is indicated in blue, cleavage site in gray, and oligonucleotides for PCR by arrows with small arrowheads.

The PCR products derived from *TOE2* cleavage products increased substantially in miR172a overexpressing seedlings (Figure 21A). At the same time, non-cleaved *TOE2* transcripts, which were amplified using oligonucleotides spanning the miRNA target site, were reduced (Figure 21A). A transient assay with *Nicotiana benthamiana* as host and *Agrobacterium tumefaciens* as delivery vehicle has been developed to study miRNA-guided cleavage (Llave et al. 2002b; Palatnik et al. 2003; Xie et al. 2003). In this system, co-transfection with a construct designed to overexpress miR172a reduced the levels of full-length *TOE2* transcript and led to accumulation of a shorter mRNA indicative of cleavage (Figure 21B). A miRNA-resistant version of *TOE2* was unaffected by miR172a, and in addition, it

accumulated to much higher levels even in the absence of exogenous miR172a (Figure 21B). Together, these experiments demonstrate that *TOE2* is efficiently guided to cleavage by miR172a, leading to much reduced steady-state levels of *TOE2* in miR172a overexpressers. We do not know, why this effect has not been detected in the activation-tagged 172 overexpresser (Aukerman and Sakai 2003).

We also detected substantial increases in the cleavage products of *AP2* and *TOE1* (Figure 21A). Since there was no corresponding decrease in *AP2* or *TOE1* steady-state levels, the simplest explanation for these findings is that miR172 targets are under direct or indirect feedback regulation. In this scenario, miR172 both causes cleavage and translational repression of its targets, and the ensuing reduced protein accumulation leads to increased transcription of targets. The action of miR172 targets as repressors of their own transcription is consistent with the known function of AP2 as a transcriptional repressor of the homeotic gene *AGAMOUS* (Bomblies et al. 1999; Deyholos and Sieburth 2000).
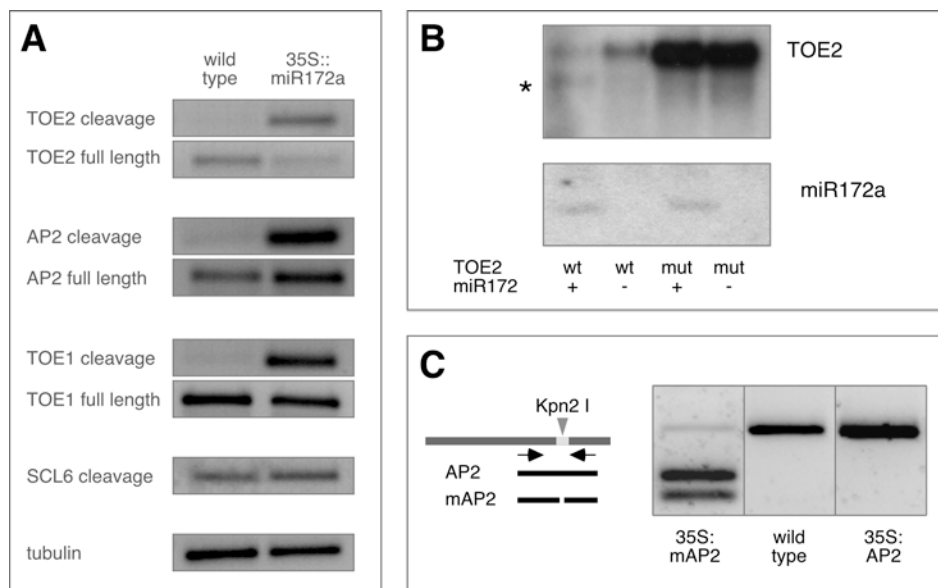
Figure 21: Cleavage and feedback regulation of miR172 targets

(**A**) Semi-quantitative analysis of uncleaved transcripts and cleavage products. Cleavage products of the *SCL6* gene (Llave et al. 2002b) and tubulin cDNAs were amplified as control. (**B**) RNA blot analysis of *N. benthamiana* leaves infiltrated with *TOE2* and miR172a overexpression constructs. Asterisk indicates cleavage product in the presence of miR172a. Transcripts from the miRNA-resistant form of *TOE2* ("mut") are more stable. (**C**) Feedback regulation of *AP2*. Wild-type and miRNA resistant ("mAP2") transcripts were amplified by RT-PCR, and distinguished by digestion with the restriction enzyme *Kpn* 2I, which cuts only the mutant form. Wild-type transcript is increased in *35S::AP2* plants, as expected, but strongly decreased in plants that presumably overproduce AP2 protein because they express a miRNA-resistant version of AP2 ("35S::mAP2").

As a test for the hypothesis of feedback regulation, we compared *AP2* RNA levels in wild-type plants and plants that overexpressed *mAP2*, a mutated, miRNA-resistant form of *AP2* (described in Chapter 1). A restriction enzyme site had been engineered into the mutated miRNA target site, allowing discrimination between wild-type and mutant *AP2* mRNA. RT-PCR demonstrated a substantial decrease in the expression of wild-type *AP2* RNA, when *mAP2* was overexpressed (Figure 21C). This experiment confirms that AP2 protein directly or indirectly represses its own transcription. We suggest that the efficiency of the feedback regulation differs for the different targets, explaining why *AP2* levels do not respond to miR172a overexpression, while the levels of other targets increase. In addition, the extent of cleavage versus translational inhibition may be affected by the number of mismatches to the miRNA, which may explain the strong miR172a-induced reduction in steady-state levels of *TOE2*, which is the only target with a single mismatch.

Similar to *AP2,* transcripts of *MYB33* and *MYB65*, targets of miR159, were also not downregulated in flowers of miR159 overexpressers. In order to test if this effect was also caused by feedback regulation, we (Javier Palatnik) quantified cleavage products for both genes, however we did not detect an obvious increase in their levels in miRNA overexpressers (not shown). This suggests again, that cleavage of *MYB33*and *MYB65* is already saturated in miR159 overexpressing plants.

Conclusions Chapter II

We have studied transcriptional changes in transgenic plants overexpressing five different miRNAs and found that most predicted target genes of miRNAs 156, 159, 164 and 319 were downregulated upon miRNA overproduction. Similar results have been observed by other studies (Achard et al. 2004; Laufs et al. 2004). These findings suggest that all four miRNAs function by cleavage-directed negative regulation of target transcript abundance, which is in line with the current model of plant miRNA action.

Since we had hybridized ATH1 microarrays with RNA from miRNA overexpressing and wild-type tissue, we could broaden our analysis to a genome wide scale and found that direct miRNA effects were mostly limited to only a few

target transcripts with high sequence complementarity to the miRNA. Targets were often related in sequence and conserved in other species. This contrasts with reports from animal miRNAs (Brennecke et al. 2005; Grun et al. 2005; Lewis et al. 2005; Lim et al. 2005) and also from application of synthetic siRNAs to animal cells (Jackson et al. 2003), where transcript with more limited complementarity to small RNAs were also affected. However, effects on target transcript levels upon animal miRNA overproduction are most likely not mediated by mRNA cleavage, but rather by destabilization of transcripts upon decapping mediated by 5'-3' exonucleases. This may reflect that natural miRNAs have co-evolved with the remainder of the transcriptome, and that there has been selection against additional targets. Alternatively, plant miRNAs or the processing machinery might have endogenous properties that make them more specific than their animal counterparts. Nevertheless, these methods did not allow us to monitor translational inhibition, which might occur on transcripts with more limited complementarity.

Comparing target sequences of confirmed targets to those of transcripts with similar mismatch numbers, but which were not downregulated in miRNA overexpressers (non-functional sequences), we determined sequence parameters, which almost unambiguously distinguish between targets and non-targets. These do not only apply to our small set of studied miRNAs, but are generally applicable with very few exceptions. Therefore, they can be used to predict target genes with high confidence when new miRNAs, especially non-conserved miRNAs, are identified.

The effects of miRNA overexpression can be complicated by feedback regulation, such that steady-state mRNA levels are not altered in the presence of excess miRNA, as we have shown for miR172. This mechanism might be used to fine-tune gene expression. Our findings suggest that transcript cleavage and translational inhibition function at the same time to regulate protein abundance of AP2-like targets and is supported by previous findings (Aukerman and Sakai 2003; Chen 2004; Kasschau et al. 2003; Lauter et al. 2005). It will be interesting to determine, to which extent translational inhibition also functions in other plant miRNA-target interactions. Experiments to shift the bias in the miR172 and other miRNA systems to either RNA cleavage or translational repression should be informative as to whether there is an advantage in having both mechanisms operate at the same time. Unfortunately, it is not yet possible to monitor genome-wide effects on protein abundance, so that individual examples of miRNA-target interactions will have to be analyzed in greater detail.

Feedback regulation appears as a general theme in miRNA biology (Baulcombe 2004). Both *DCL1* and *AGO1*, two central components of miRNA

biosynthesis and action are themselves under miRNA regulation (Vaucheret et al. 2004; Xie et al. 2003). In addition, miR159 and its target *MYB33* are both induced by a common upstream stimulus (Achard et al. 2004).

CHAPTER III

ARTIFICIAL microRNAS

This work has been submitted as a manuscript to Plant Cell

Artificial microRNAs Silence Target Genes with High Specificity

Rebecca Schwab, Stephan Ossowski, Markus Riester, Norman Warthmann, and Detlef Weigel

Contributions to this chapter:

RS designed and carried out all experiments and analyses apart from generating amiR-ft overexpressing plants

SO developed the algorithm for automated amiRNA design

MR generated the web-tool for automated amiRNA design.

NW generated amiR-ft overxpressing plants

DW had the initial idea to design amiRNAs, suggested experiments and discussed results.

Most of this chapter is adapted from the manuscript mentioned above, which has been written by RS and DW.

Plant miRNAs negatively regulate small numbers of target genes, to which they share high sequence complementarity, by cleavage induced transcript degradation (Kasschau et al. 2003; Mallory et al. 2004b; Palatnik et al. 2003; Schwab et al. 2005). In animals, only limited sequence complementarity is sufficient for target recognition, so that the typical animal miRNA regulates more than a hundred targets upon pairing to the so-called seed region in the 5' part of animal miRNAs, which induces inhibition of target translation (Brennecke et al. 2005; Farh et al. 2005; Lewis et al. 2005; Lim et al. 2005). SiRNAs function by transcript cleavage, similar to plant miRNAs, and perfectly complementary siRNAs are widely used in animals as a tool to down-regulate RNA expression of genes of interest. They can either be synthesized in vitro or by transgenic expression of a double stranded precursor with hairpin structure (Hannon and Rossi 2004). Sequence parameters that lead to particularly effective gene silencing by siRNAs have been identified though systematic analyses of siRNA effects (Reynolds et al. 2004). However, siRNAs can also affect RNAs that are not perfectly complementary, generally considered off-targets (Doench and Sharp 2004; Haley and Zamore 2004; Jackson et al. 2003). The differences in specificity between plant and animal small RNA action have raised the question weather these are only due to intrinsic properties of the small RNA machineries or if they are at least partially caused by selection against plant miRNAs with large numbers of targets. To address this problem, I generated artificial miRNAs (amiRNAs) targeting endogenous mRNAs and compared their effects to those of natural plant miRNAs.

Design of artificial microRNAs

The analysis of plants that overexpress natural miRNAs, together with a reexamination of known targets, had led us to propose specific sequence parameters important for target selection by plant miRNAs (Schwab et al. 2005). We had found that pairing to the 5' portion of the miRNA (positions 2 to 12) was most important, since this region often had no mismatch and rarely more than one. Similarly, mismatches at the presumptive cleavage site (positions 10/11) were usually not present in direct targets. In vitro experiments with mutant targets largely support these findings (Mallory et al. 2004b). Clusters of more than two mismatches to the 3' part of the miRNA were rare. In addition, perfect pairing in the 3' portion can compensate for the presence of up to two mismatches in the 5' portion, leading to a low overall free energy of targets paired with their corresponding miRNAs (at least 70% as compared to a perfect match and a maximum of –30kcal/mole).

The same parameters were incorporated into the design of amiRNAs. I began by selecting different target genes, most of which had known loss-of-function phenotypes that could be easily monitored. In addition, the amiRNAs were designed with uridine at position 1 and adenine at position 10, both of which are over-represented among natural plant miRNAs and highly efficient siRNAs (Mallory et al. 2004b; Reynolds et al. 2004). I also preferred amiRNAs to display 5' instability relative to their miRNA*, so that the correct sequence would be recruited to AGO1. AmiRNA sequences fulfilling the functionality criteria were initially selected by hand from reverse complements of target genes. To reduce the likelihood that an amiRNA would act as primer for RNA dependent RNA polymerases, and thereby trigger secondary RNAi, between one and three mismatches to the target genes were introduced in the 3' part of the amiRNAs. Alignments of amiRNAs and all intended targets are listed in Supplementary Fig. 2.

It has been previously shown that both animal and plant miRNA precursors can be modified to express a small RNA with a sequence that is unrelated to the miRNA normally produced by the precursor (Parizotto et al. 2004; Zeng et al. 2002). I used precursors for miRNA172a and miR319a as backbones for amiRNA expression under control of the constitutive 35S promoter from cauliflower mosaic virus. Using overlapping PCR, I exchanged the natural miRNA sequences with those of amiRNAs. I also modified the miRNA* region, which base-pairs to the miRNA in the precursor, such that both structural and energetic features of the miRNA precursor were retained (Figure 22).

Phenotypic as well as molecular analysis of amiRNA plants was carried out in primary transformants (T1 generation). The effects of amiRNA overexpression could, however, be stably inherited, as seen with an amiRNA directed against the flowering time gene *FT* (not shown). A list of amiRNAs and their targets is provided in Table 14.
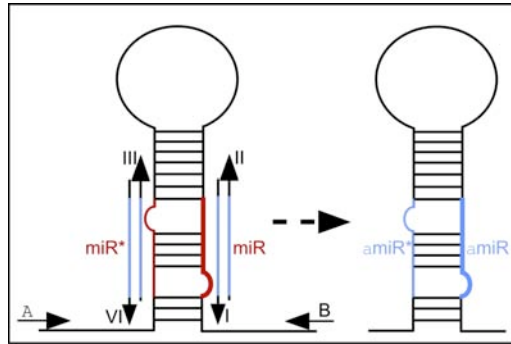
Figure 22: Engineering artificial miRNAs

Site-directed mutagenesis on precursors of endogenous miRNAs was carried out using overlapping PCR. Oligonucleotide primers I to IV were used to replace miRNA and miRNA* regions (red) with artificial sequences (blue). Primers A and B were based on template plasmid sequence. Regeneration of functional miRNA precursors was achieved by combining PCR products A-IV, II-III and I-B in a single reaction with primers A and B.

Table 14: Predicted amiRNA sequences and targets

amiRNA	Predicted mature sequence	Predicted target(s)	Known target functions	Ref.
amiR-lfy-1	**U**AACAGUGA**A**CGUACUGUCGC	*LFY*	master regulator of floral identity	1
amiR-lfy-2	**U**UACGAUA**A**CGGUUGCUCGC			
amiR-white-1	**U**UAGUGAGA**A**UGUUGCGCCGG	*GUN4*	cofactor in chlorophyll biosynthesis	2
amiR-white-2	**U**UUAACCAGA**U**UUUGCGUCGC			
amiR-ft-1	**U**AUUCUCGG**A**GGUGAGUGUUG	*FT*	promotion of flowering	3
amiR-ft-2	**U**UGGUUAUA**A**AGGAAGAGGCC			
amiR-trichome	**U**CCCAUUCG**A**UACUGCUCGCC	*TRY, CPC, ETC2*	trichome patterning	4,5
amiR-mads-1	**U**UUUGGAGA**A**AGUGACUUGUC	*SOC1, MAF1-3, ANR1,* and 3 others	regulation of flowering, nutrient uptake	6
amiR-mads-2	**U**UGUUCUCU**A**UCCUCUUCAGC	*SEP1-4, SHP1-2, AP1, CAL* and 10 others	patterning of floral organs	
amiR-yabby-1	**U**ACUGAAAG**C**UUCUCUGUGGG	*INO, YAB3,* and 3 others	regulation of adaxial polarity	7
amiR-yabby-2	**U**GUAUGCUG**A**UGGGACUCUCG	*CRC*		

References: 1, (Weigel et al. 1992); 2, (Larkin et al. 2003); 3, (Kardailsky et al. 1999); 4, (Schellmann et al. 2002); 5, (Kirik et al. 2004); 6, (Becker and Theissen 2003); 7, (Engstrom et al. 2004). Full list of targets in Supplementary Table 1.

Molecular identity of artificial microRNAs

To confirm that amiRNAs accumulated in transgenic plants, I hybridized small RNA blots of inflorescence tissue from pooled T1 plants. All amiRNAs tested were efficiently expressed from both *MIR319a* and *MIR172a* backbones (Figure 23). Differences in the mobility of the amiRNAs on polyacrylamide gels might reflect either some heterogeneity in size, or might be due to sequence differences, as shown with mutant forms of miR159a (Palatnik et al., manuscript submitted). This suggests that the majority of artificial miRNAs were 21 nucleotides in length, as intended. In some cases, small RNAs of different length accumulated, indicating that the position of DICER-LIKE1 cleavage was not uniform in all cases. Only very weak or no signals were detected with miRNA* specific probes, with the exception of the miRNA* for amiR-trichome (not shown). amiR-trichome was the only amiRNA without 5' instability relative to its miRNA*, indicating that selection of the amiRNA from the double-stranded DICER-LIKE1 product was similar to siRNA strand selection.
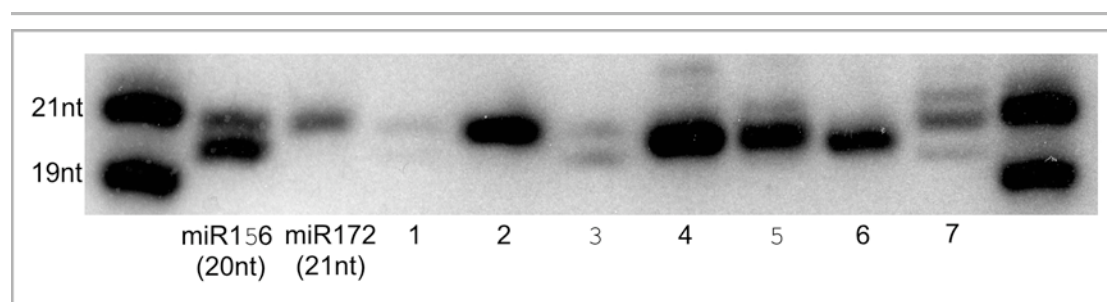
Figure 23: Expression of amiRNAs

RNA blot analysis of amiRNA overexpressers using a mixture of probes for all amiRNAs. The outermost lanes contain two standards. miR156a and miR172a overexpressers were included as control. 1, amiR-mads-2 (*MIR319a* backbone); 2, amiR-mads-1 (*MIR172a* backbone); 3, amiR-trichome (*MIR319a* backbone); 4, amiR-lfy-2 (*MIR319a* backbone); 5, amiR-lfy-2 (*MIR172a* backbone); 6, amiR-lfy-1 (*MIR172a* backbone); 7, amiR-yabby-2 (*MIR319a* backbone).

Because AGO proteins cleave targets invariably opposite of position 10/11 of the small RNA (Kasschau et al. 2003), the 5' ends of small RNAs can be inferred by mapping the cleavage products of their targets (Llave et al. 2002b). For targets of amiR-mads-1 (*MIR172a* backbone), amiR-mads-2 (*MIR319a* backbone), and amiR-trichome (*MIR319a* backbone), cleavage products had the expected 5' ends (Figure 24). Since uniform cleavage products were obtained even in cases where there was

some size heterogeneity in the amiRNA, suggests that these amiRNAs differed only at their 3' end.

AmiR-lfy-1 (*MIR172a* backbone) caused cleavage of the target two nucleotides downstream of the expected position, implying that the initial DICER-LIKE1 product was shifted by two nucleotides. Examination of the sequence surrounding the intended amiRNA in the precursor revealed that this alternative amiRNA would still be specific for its target (Figure 24).

Taken together, artificial miRNAs were effectively produced from their precursors, and a high fraction was processed as the exact 21mer that was exchanged in the backbone precursor.
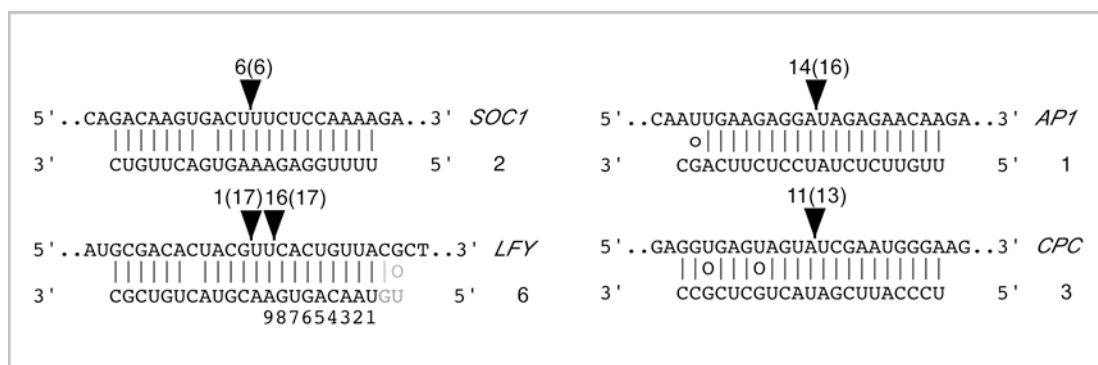
Figure 24: Cleavage of predicted amiRNA targets

Mapping of target cleavage products by RACE-PCR. Parentheses indicate total number of clones analyzed. In the case of *LFY*, only one clone had a 5' end at the expected position, opposite nucleotides 10 to 11 of the intended amiRNA. The 5' end of most clones was offset by two nucleotides, suggesting that most of the amiRNAs were offset as well. The sequence predicted from the aberrant processing is indicated in grey.

Effects on predicted target genes

Single targets. Overexpression of three amiRNAs designed to target single genes resulted in robust and strong phenotypes that resembled those of plants with mutations in the respective target gene (Figure 25, Table 18). In most cases, more than 90% of primary transformants displayed defects, but the fraction of plants that resembled null mutants varied depending on amiRNA transgenes and precursor
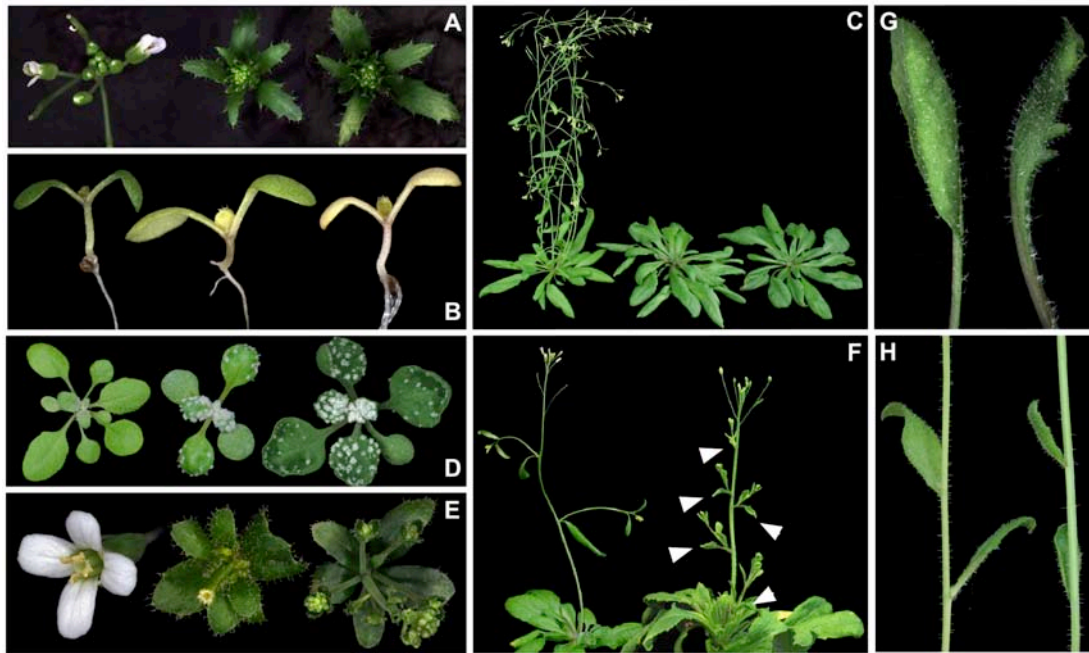
Figure 25: Phenotypes of amiRNA overexpressers

(A) Inflorescences. From left to right, wild type, *lfy*-12, and amiR-lfy-1 (*MIR172a* backbone) overexpresser. **(B)** Seedlings. From left to right, wild type, *gun4*-1, amiR-white-1 (*MIR172a* backbone) overexpresser. Bleaching of cotyledons is more pronounced in the amiR-white plants than in *gun4*-1, consistent with the more severe molecular profile of the amiR-white overexpressers. **(C)** Adult plants sown on the same day. From left to right, wild type, *ft*-10, and amiR-ft-2 (*MIR172a* backbone). **(D)** Leaf rosettes. From left to right, wild type, *try cpc* double mutants, and amiR-trichome (*MIR319a* backbone) overexpresser. Clustered trichomes are evident even at low magnification. **(E)** Flowers. From left to right, wild type, weak amiR-mads-2 (*MIR319a* backbone) overexpresser, and strong amiR-mads-2 (*MIR319a* backbone) overexpresser. In both amiR-mads overexpressers, outer whorls are transformed into leaf-like structures. In the strong line, secondary inflorescences replace the central gynoecium. **(F)** Flowering plants. Left, wild type; right amiR-mads-1 (*MIR172a* backbone) overexpresser with increased number of cauline leaves (arrowheads). **(G)** Rosette leaves of wild type (left) and amiR-yabby-1 (*MIR172a* backbone) overexpressers. Abaxial side is left. **(H)** Cauline leaves of wild type (left) and amiR-yabby-2 (*MIR319a* backbone) overexpressers (right) with polarity defects.

---

backbones (see below). The majority of amiR-lfy-1 overexpressers had floral defects resembling *lfy* null mutants (Figure 25A), while others showed milder effects more typical of weak and intermediate *lfy* alleles (Weigel et al. 1992). Similarly, most amiR-white overexpressers were arrested in their growth as white seedlings, similar to *gun4* mutants (Figure 25B) (Larkin et al. 2003). The most consistent effects were seen with amiRNAs targeted against the flowering time gene *FT*, whose loss of function results in late flowering under long days (Koornneef et al. 1991). All plants overexpressing either amiR-ft-1 or amiR-ft-2 (*n* = 40) flowered within a day of *ft* null mutants (Figure 25C) (transgenic plants generated by Norman Warthmann).

I used RT-PCR and microarray analyses to examine the effects of amiRNAs on their targets. *FT* transcripts were decreased below detection level by RT-PCR, similar to *ft* T-DNA insertion mutants (Figure 26). *LFY* and *GUN4* transcripts were still detectable in amiR-lfy-1 inflorescences and amiR-white-1 seedlings, respectively, using Affymetrix arrays, but were substantially reduced (4.5 fold, and 5.7 fold). *GUN4* was no longer detectable in amiR-white-2 overexpressers using Affymetrix arrays.
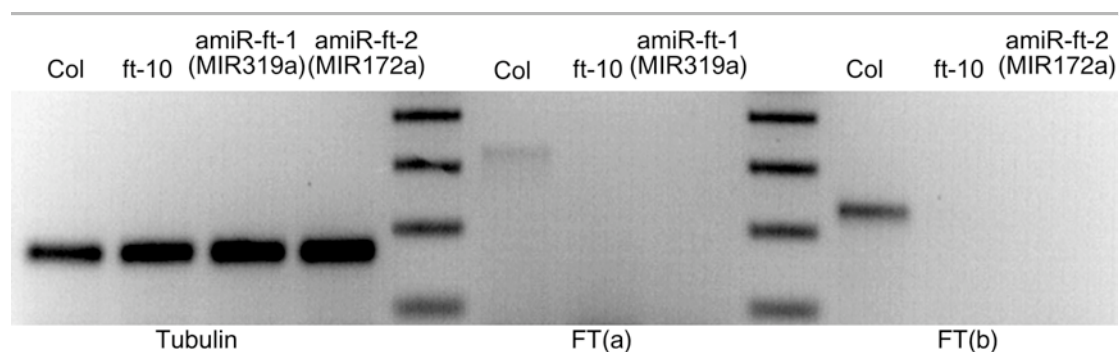
Figure 26: Expression of *FT* in leaves of amiR-ft overexpressers

RT-PCR reactions were stopped in the linear phase of amplification.

Multiple targets. Because many natural miRNAs have several targets, I designed three classes of amiRNAs with multiple potential targets from different transcription factor gene families. amiR-trichome targets three *MYB* genes, *CAPRICE* (*CPC*), *TRIPTYCHON* (*TRY*) and *ENHANCER OF TRIPTYCHON AND CAPRICE 2* (*ETC2*), which are involved in trichome patterning (Kirik et al. 2004; Schellmann et al. 2002). amiR-mads-1 and amiR-mads-2 target seven and 13 MADS box genes, respectively, that are mostly involved in the control of flowering time and floral patterning (Becker and Theissen 2003), with a few additional non-MADS targets (complete list in Supplementary Table 1). AmiR-yabby-1 and amiR-yabby-2 target different members of the *YABBY* gene family, with two predicted targets for amiR-yabby-1 and one for amiR-yabby-2. *YABBY* genes specify the abaxial domain of lateral organs (Engstrom et al. 2004). The target motif for amiR-trichome was located outside the region encoding the DNA binding domain, which is the typical case for natural miRNAs targeting transcription factor genes, while the target motifs for amiR-mads and amiR-yabby were in the regions encoding the DNA binding domains.

In most cases, the phenotypes of amiRNA overexpressers suggested in vivo targeting of multiple genes. About 30% of amiR-trichome overexpressers

82

phenocopied *try cpc* double mutants with highly clustered trichomes on leaf blades (Schellmann et al. 2002) (Figure 25D), while most of the remaining amiR-trichome overexpressers resembled *try* or *cpc* single mutants. A third potential target, *ETC2*, predicted to be a much less favorable target than *TRY* or *CPC*, seemed to be affected in only very few plants, which had extra trichome clusters on petioles, as seen in *try cpc etc2* triple mutants (Kirik et al. 2004) (not shown).

Potential targets of amiR-mads-1 include both floral repressors such as the *MAF* genes (Ratcliff et al. 1999; Scortecci et al. 2001), and floral activators such as *SOC1* (Borner et al. 2000; Lee et al. 2000; Samach et al. 2000), as well as *ANR1*, which is involved in nutrient uptake (Zhang and Forde 1998). Even though chronological flowering time was not greatly altered in overexpressing plants, cauline leaf number was increased from two or three to four or five in >90% of T1 plants grown in continuous light (Figure 25F). In addition, there were carpel defects that have not been described for any of the intended targets, although there are several MADS box genes such as *AG*, *FUL* and *SHP1/2* that are involved in fruit patterning (Becker and Theissen 2003). Microarray analysis of inflorescences of amiR-mads-1 overexpressers showed down-regulation of the four intended targets that were present in the control, although the effects were small, between 1.4 and 2.3 fold (Figure 27C), and not statistically significant, using the logit-T algorithm at $p < 0.025$ (Lemon et al. 2003).

The targets predicted for amiR-mads-2 are mostly required for determining floral organ identity (Becker and Theissen 2003). Flowers of overexpressers had severe changes in floral morphology with leaf-like organs in all four whorls, characteristic of *sepallata* multiple mutants (Pelaz et al. 2000) (Figure 25E). Defects were stronger in the two outer whorls compared to the two central whorls, which might be due to non-uniform activity of the cauliflower mosaic virus 35S promoter I used. Additional features like secondary flowers, indeterminacy of floral meristems, and incomplete separation of carpel valves have been described as features of other predicted target gene mutants such as *ap1* and *shp1/2* (Irish and Sussex 1990; Liljegren et al. 2000)*.* In the strongest lines, carpels were replaced by a new inflorescence shoot. Expression analysis with Affymetrix microarrays showed that most predicted target genes were significantly down-regulated (logit-T, $p < 0.025$) in inflorescence tissue of both weak and strong overexpressers (Figure 27B and Supplementary Table 1).
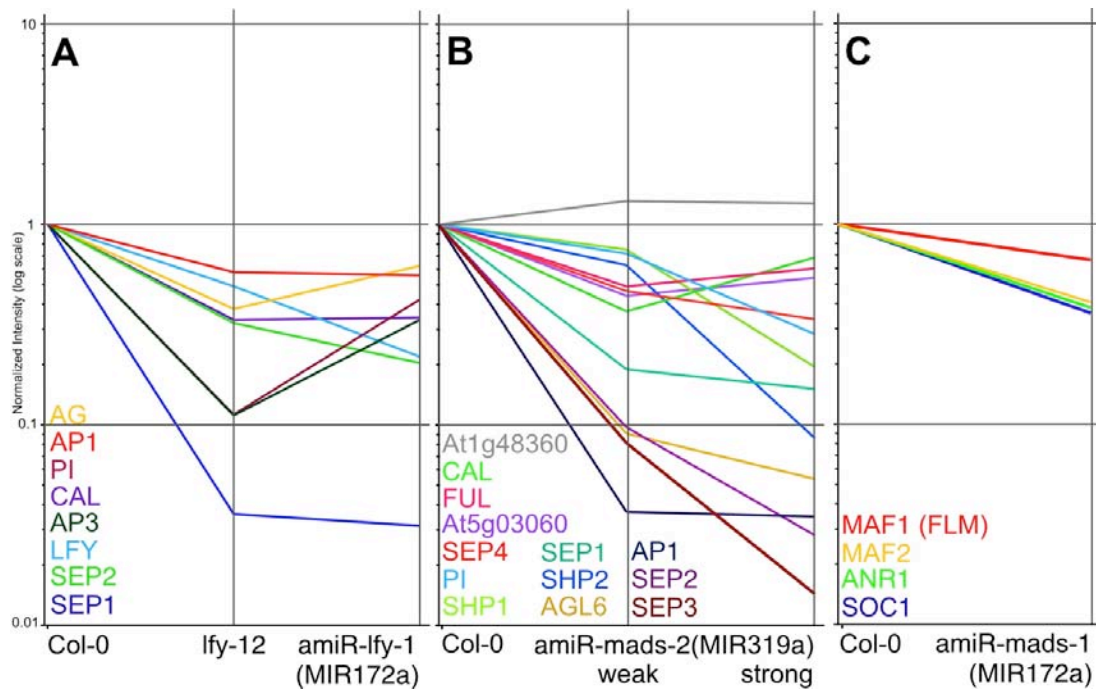
Figure 27: Expression analysis of amiRNA overexpressers

(A) Microarray profiles of *LFY* and some of its direct downstream targets in inflorescences of wild type (Col-0), *lfy*-12 mutants, and amiR-lfy-1 (*MIR172a* backbone) overexpressers. **(B)** Microarray profiles of predicted amiR-mads-2 targets in inflorescences of wild type (Col-0), and weak and strong amiR-mads-2 (*MIR319a* backbone) overexpressers. **(C)** Microarray profiles of predicted amiR-mads-1 targets in wild type (Col-0) and amiR-mads-1 (*MIR172a* backbone) overexpressing inflorescences.

Plants overproducing amiR-yabby-1 and -2 had defects in leaf polarity, such as leaves with trichomes on both sides, indicating adaxialization (amiR-yabby-1, Figure 25G). Other phenotypes, such as polarity defects in cauline leaves (amiR-yabby-1 and -2, Figure 25H*)*, while not described for *yabby* mutant combinations before, are likely also related to *YABBY* function in polarity establishment. RT-PCR analyses with RNA from amiR-yabby-1 overexpressing seedlings and inflorescences showed predicted targets to be downregulated. Surprisingly, expression levels of the target gene *CRC* remained unchanged in amiR-yabby-2 (Figure 28). When I examined other *YABBYs* as potential targets of amiR-yabby-2, I found that *FIL* was down-regulated, and mapping of cleavage sites confirmed that *FIL* was targeted by amiR-yabby-2 (not shown). FIL was initially not considered as a target, because there are two mismatches in the critical 5' region of the amiRNA (positions 2 to 12). However, the mismatches are at positions 2 and 8, with only one additional mismatch at the 3' end of the miRNA, which makes this interaction similar to the one observed for miR168 and its target *AGO1* (Rhoades et al. 2002; Vaucheret et al. 2004). This
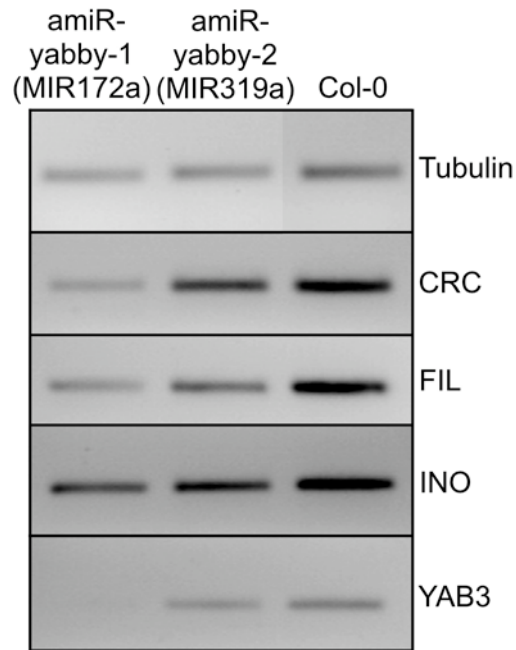
Figure 28: Expression of *YABBY* genes in inflorescence tissue of amiR-yabby overexpressers

RT-PCR reactions were stopped in the linear phase of amplification.

pair was one of the few exceptions that we had found to our more restrictive general rules that hold for the vast majority of natural miRNA targets (Schwab et al. 2005).

In summary, multiple mRNAs can be successfully targeted by amiRNAs. While the degree of downregulation varied for different targets, there was no clear correlation of targeting efficiency either with the extent of complementarity between amiRNA and target, or with expression levels of targets in wild type. Thus, accessibility of target sites or feedback regulation of target transcripts might play additional roles. Since the related targets had the amiRNA complementary motif at approximately the same position in the coding region, such accessibility differences would not be due to the relative position within the transcript. In support of this, both amiRNAs directed against the beginning of the coding region (+125 for amiR-white-1) and ones directed against the 3' UTR (amiR-ft-2) were effective.

Effect of backbones. Comparing the effectiveness of the different stem loop backbones used, I conclude that both *MIR319a* and *MIR172a* precursors can be used for amiRNA expression. However, more robust results were obtained with *MIR319a* derivatives, all of which led to phenotypic changes, while this was not the case of *MIR172a*. For several targets, more than one amiRNA has been tested

(Table 18). In general, they caused similar defects, but they were not always equally effective. For example, amiR-lfy-2 caused *lfy*-like phenotypes only when expressed from *MIR319a* backbone, while amiR-lfy-1 caused strong *lfy* defects also when expressed from the *MIR172a* backbone.

Specificity and non-transitivity of artificial microRNAs

Using microarray analyses, we have previously determined parameters for target selection by natural *Arabidopsis* miRNAs, and found that plant miRNAs are apparently much more specific than animal miRNAs (Brennecke et al. 2005; Lewis et al. 2005; Lim et al. 2005; Schwab et al. 2005). I used a similar approach to investigate the specificity of amiRNAs. Globally, I found that the significantly down-regulated genes (determined by logit-T probe testing at $p < 0.025$) in each amiRNA overexpresser were on average not more similar to the respective amiRNAs than all genes, once predicted targets had been removed (Figure 29, Table 15).
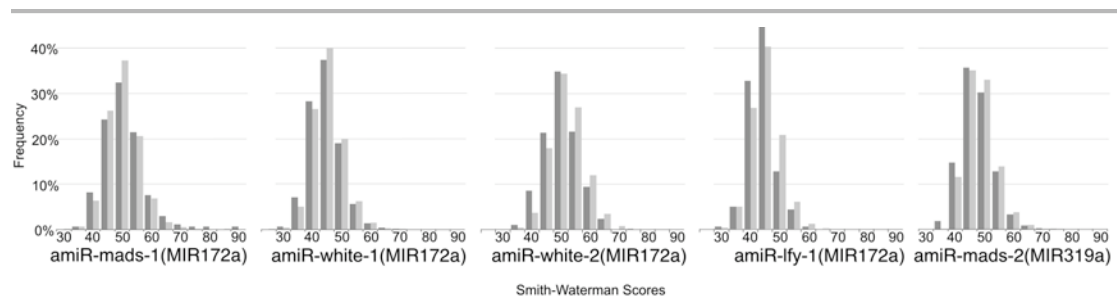
Figure 29: Specificity of amiRNAs

Distributions of Smith-Waterman scores (Smith and Waterman 1981) are similar between genes that are significantly downregulated in response to amiR overexpression (dark grey bars) and all genes present in the control (light grey bars). Predicted targets have been removed.

Natural miRNAs have targets with up to five mismatches, although the large majority of these genes are not targets. I therefore focused more specifically on mRNAs with up to five mismatches to the different amiRNAs, assuming that all direct target genes would be found among this group. If amiRNAs were indeed specific for our intended target genes, then the fraction of genes with up to five mismatches should not be overrepresented among downregulated genes, once predicted targets have been removed.

Table 15: Means of Smith-Waterman scores

Experiment	Control mean	Experimental mean	t	p-value
35S::amiR-lfy-1(*MIR172a*)	47.2	46.2	3.902	0.9999
35S::amiR-white-1(*MIR172a*)	47.2	46.8	2.589	0.9950
35S::amiR-white-2(*MIR172a*)	50.3	49.8	3.790	1.0000
35S::amiR-mads-1(*MIR172a*)	52.2	53.1	-1.693	0.9530
35S::amiR-mads-2(*MIR319a*)	54.1	52.6	7.283	1.0000

All genes present in the control (wild-type Col-0) were compared with the genes significantly downregulated in response to amiRNA overexpression (experiment). One-sided student's t-tests suggest that the means of the experimental samples are not significantly higher than those of the controls. (In most cases, they are actually lower.)

To err on the safe side, used a per-gene change (logit-T, $p < 0.025$) as well as a common change (1.5 fold reduction). In no case was the difference between observed and downregulated "non-targets" with up to five mismatches statistically significant (Table 16), although amiR-mads-1 downregulated substantially more genes than expected. Most of these were MADS box genes, for which extensive cross-regulation during flowering and floral patterning is well known (Becker and Theissen 2003). RACE-PCR was carried out for the important floral repressor, *FLC*, a MADS box gene that had not been predicted as a target because it has a mismatch to position 11 of amiR-mads-1. No cleavage products were detected, suggesting that the overrepresentation of MADS box genes among downregulated genes was due to secondary effects.

As a third measure for artificial miRNA specificity, I compared genome-wide expression profiles of amiRNA overexpressers targeting single genes (*LFY* and *GUN4*) with those of plants with mutations in the target genes. The majority of genes downregulated in amiR-lfy-1 plants was also affected in *lfy*-12 mutants, and included known direct downstream targets of *LFY* itself (Figure 27A, Figure 30A).

Both amiR-white-1 and amiR-white-2, which target different regions of the *GUN4* mRNA and are unrelated in sequence, caused downregulation of very similar sets of genes (Figure 30B). The number of differentially expressed genes was, however, much higher than in *gun4*-1 mutants, which can most likely be attributed to severity of phenotypes, since the mutation in *gun4*-1 merely causes an amino acid replacement, and the *gun4*-1 allele is likely hypomorphic (Larkin et al. 2003)

Table 16: Summary of downregulated genes with up to 5 mismatches

		amiR-lfy-1	amiR-mads-1	amiR-mads-2	amiR-white-1	amiR-white-2
Present in control		32	89	311	16	29
Observed downregulated	Predicted targets	1 (1)	3 (0)	11 (8)	1 (1)	1 (1)
	Others	1 (1)	5 (4)	25 (13)	7 (3)	7 (3)
Expected downregulated without predicted targets		1.5 (0.6)	1.8 (0.9)	29.9 (18.5)	4.3 (2.1)	9.3 (4.3)

Downregulated genes among all genes with up to 5 mismatches to the respective amiRNA, using a common change (1.5 fold) or per-gene change (logit-T, $p < 0.025$) in parentheses. In no case was the difference between expected and observed down-regulated genes among the "non-targets" statistically significant (Fisher's Exact or χ^2 test). 15,368 genes were called present in Col-0 inflorescences (control to amiR-lfy-1, amiR-mads-1, and amiR-mads-2), 14,039 genes in Col-0 seedlings (control to amiR-white-1 and amiR-white-2). For amiR-lfy-1, the sequence of the 21mer deduced from the mapping of the cleavage product was used, which was shifted by 2 nucleotides from the intended amiRNA.
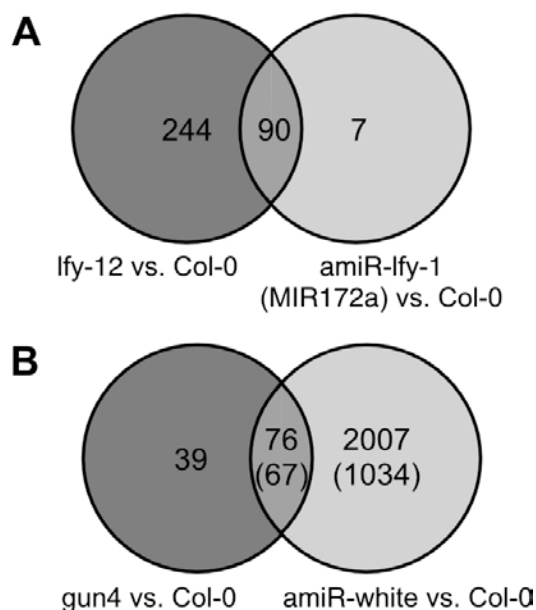
Figure 30: Significantly downregulated genes on microarrays

(A) Overlap of significantly downregulated genes (logit-T, $p < 0.01$) in *lfy*-12 and amiR-lfy-1 plants indicates a very similar molecular phenotype, with amiR-lfy-1 plants being on average weaker than *lfy*-12 plants. **(B)** Overlap of significantly changed genes in *gun4*-1 and amiR-white overexpressers. Numbers in parentheses indicate genes common to amiR-white-1 (*MIR172a* backbone) and amiR-white-2 (*MIR172a* backbone) overexpressers. Only genes called as present by Affymetrix algorithms in wild-type controls were considered.

Finally, I examined whether amiRNAs are likely to have indirect effects through a process called transitivity. Upon binding to target transcripts, siRNAs cannot only trigger their cleavage and subsequent destruction, but also serve as primers for RNA dependent RNA polymerases. These extend the local RNA double strands and generate templates for production of secondary siRNAs by Dicer action (Voinnet 2005). These secondary siRNAs, which are unrelated in sequence to the initial trigger, can in turn affect other genes not targeted by the original small RNA. For two *Arabidopsis* miRNAs, miR173 and miR390, which both bind non-coding RNAs as primary targets, similar mechanisms have been described (Allen et al. 2005). To investigate the possibility of transitivity, I determined potential 21mer secondary siRNAs for all artificial miRNA target genes from both strands of a 250 bp region, surrounding the initial binding site of the amiRNA, as described by Allen and colleagues (2005). Potential targets of these siRNAs were identified using our miRNA:target algorithms. Examination of microarray data did not reveal any evidence for effects on such secondary targets (Table 17). In summary, all data suggest that the specificity of amiRNAs is very similar to that of natural miRNAs.

Table 17: Potential targets of secondary siRNAs

	amiR-lfy-1	amiR-white-1	amiR-white-2	amiR-mads-1	amiR-mads-2
Number of considered 21mers	13	15	14	61	153
Number of potential targets (no direct targets)	6	69	13	114	270
Observed downregulated of potential targets	0	2	0	4	9
Expected downregulated without direct targets	0.1	6.2	1.4	1.3	12.2

Potential secondary siRNAs for five amiRNAs were predicted according to (Allen E, Xie Z, Gustafson AM, Carrington JC (2005) microRNA-directed phasing during trans-acting siRNA biogenesis in plants. Cell 121: 207-221). Of 22 21mers from a 250 bp region surrounding the original amiRNA complementary site of each target gene, those displaying 5' instability or symmetrical thermodynamic stability relative to their pairing strand were chosen. miRNA target determinants were used to predict targets of these 21mers and significant expression changes were determined by logit-T probe testing ($p < 0.025$) on microarray data.

Temporally and spatially restricted expression of artificial microRNAs

The exquisite specificity of amiRNAs suggested that they constitute an excellent gene silencing tool, because of the predictability of their effects, especially when targeting multiple genes. To further explore the usefulness of amiRNAs, I first asked whether it is possible to transiently knock down gene expression, which has recently been demonstrated for conventional hairpin RNAi constructs as well (Wielopolska et al., 2005). Using an inducible expression system based on the ethanol-responsive *Alc* regulon (Roslan et al. 2001), both amiR-white-1 and amiR-trichome produced the expected phenotypes within three days of ethanol application. Importantly, the effects were transient (Fig. 31A, B), suggesting that the amiRNAs do not have secondary effects due to RNAi or DNA/chromatin modification, which can be transmitted autonomously after an initial triggering event.

Next, I asked whether the effects of amiRNAs can be spatially restricted by expressing them under the control of tissue-specific promoters, similar to what has been shown for RNAi using hairpin constructs (Byzova et al. 2004). amiR-lfy-1 was expressed from the *LFY* promoter (Blazquez et al. 1997), and found to result in plants resembling *lfy* mutants (Figure 31D). amiR-mads-2, which is predicted to target several MADS box homeotic genes and amiR-white-1 targeting *GUN4*, were expressed throughout the early flower and later in the outer two whorls using the *AP1* promoter (Hempel et al. 1997). AP1:amiR-white-1 produced pale inflorescences (Figure 31C) and strong AP1:amiR-mads-2 lines resembled *ap1 cal* double mutants (Bowman et al. 1993), while weaker lines were more similar to *ap1* single mutants (Figure 31C,D). The meristem identity defects of strong lines were more severe than those of 35S:amiR-mads-2 plants (Figure 25E). amiR-mads-2 was also expressed in the inner two floral whorls using regulatory elements located in the second intron of *AG* (Busch et al. 1999). The effects of this construct were not quite as severe as seen in the strongest 35S:amiR-mads-2 plants (Figure 31D). The whorl-specific effects of amiR-mads-2 when placed under the control of *AP1* and *AG* regulatory sequences indicate that there are no long-range effects. However, effects of amiRNAs do not seem to be completely cell-autonomous, since pale yellow seedlings were obtained when amiR-white-1 was expressed from the epidermis-specific *ML1* promoter (Sessions et al. 1999) (Figure 31E). Since chloroplasts are restricted to the sub-epidermal mesophyll cells this amiRNA can most likely move across at least one cell boundary. The effects were, however, much milder than with the 35S:amiR-white-1 transgene, which led to growth arrest of seedlings devoid of chlorophyll, similar to *gun4* null mutants (Figure 25B). In addition, leaf margins for

90

ML1:amiR-white-1 plants were paler than the central part of the leaves, consistent with limited movement, since the margins contain fewer cell layers (Figure 31E).
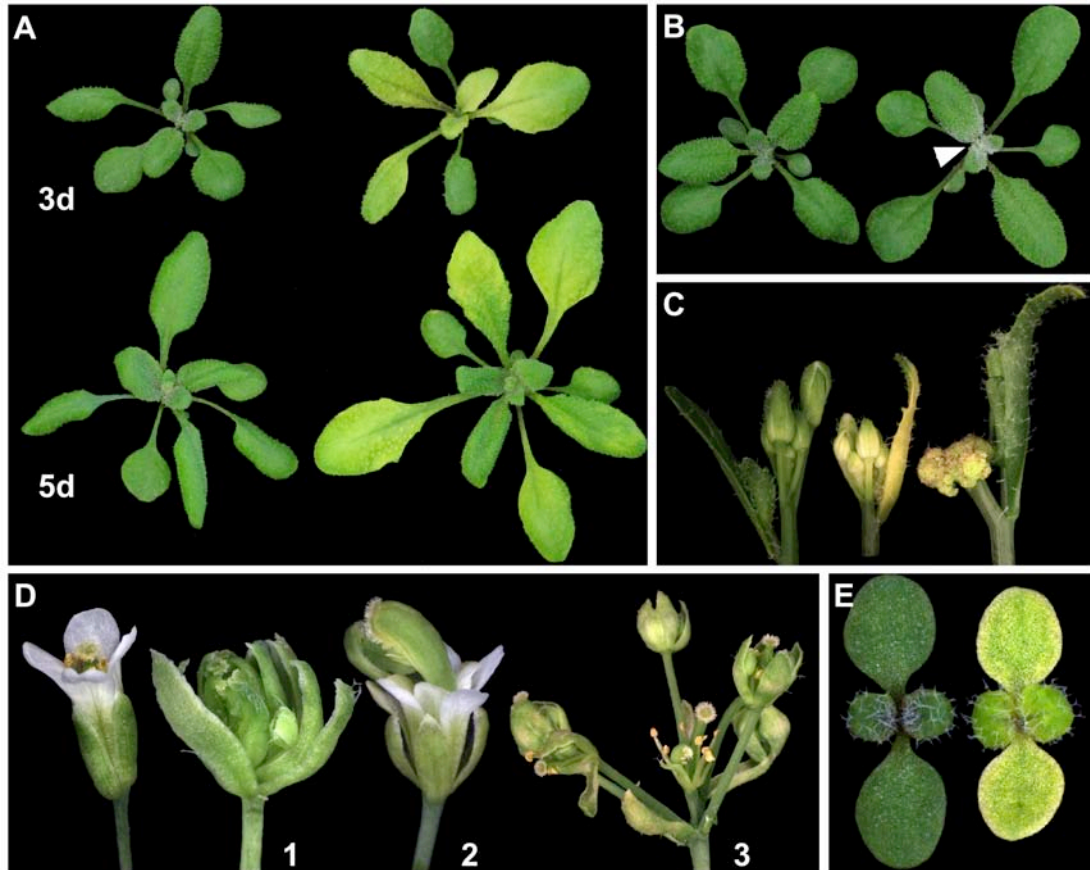
Figure 31: Inducible and tissue-specific expression of amiRNAs

Un-induced or wild-type controls are always shown on the left. **(A)** Ethanol-induced ubiquitous expression of amiR-white-1 3 and 5 days after induction. After 3 days, young leaves are all yellow; after 5 days, the youngest leaves are again green. **(B)** Ethanol-induced ubiquitous expression of amiR-trichome (right), 3 days after induction. Clustered trichomes appear as white covering of youngest leaves (arrowhead). **(C)** Inflorescences of plants expressing amiR-white-1 from the *AP1* promoter (middle) are pale yellow. Strong lines expressing amiR-mads-2 from the *AP1* promoter (right) resemble *ap1 cal* double mutants. **(D)** Expression of amiR-lfy-1 from the *LFY* promoter (1) results in flowers resembling *lfy* mutants. amiR-mads-2 expressed from *AG* regulatory elements in the center of the flower (2) produces organ transformations in the central two whorls (3). Outer whorls remain unaffected. An opposite phenotype is seen after expression of amiR-mads-2 from the *AP1* promoter (4, weaker line), which does not affect inner whorls, but results in secondary flowers, resembling *ap1* mutants. **(E)** Epidermal expression of amiR-white-1 from the *ML1* promoter results in pale plants.

Automated design of amiRNAs with the WMD tool

(by Stephan Ossowski and Markus Riester)

To facilitate the application of the amiRNA technology, we have developed a web-based tool for their automated design (Web MicroRNA Designer, WMD). The program uses sequences of target genes as an input and searches for candidate 21mers that resemble natural miRNAs in reverse complements of these genes, using the criteria described above. Target genes in the *A. thaliana* genome are determined for individual candidates using a HyPa/vmatch search tool, which is based on a suffix array algorithm to identify sequence patterns (Graf et al. 2001) and subsequent filtering according to rules for miRNA targeting based on our earlier work (Schwab et al. 2005). Mismatches in the 3' part of candidate sequences are used to reduce the possibility of off-target effects. Oligonucleotide sequences for generation of amiRNA precursors through overlapping PCR are generated as a final output. Input sequences are not restricted to *A. thaliana* sequences, so that amiRNAs that target orthologous genes from different species can be easily designed. The web-tool will be available at http://wmd.weigelword.org.

Conclusions chapter III

In contrast to miRNAs in animals, natural plant miRNAs have a very narrow action spectrum and target only mRNAs with few mismatches. I have overexpressed different artificial miRNAs (amiRNAs) in *Arabidopsis thaliana* and found that similar parameters of target selection apply as for natural miRNAs, and that direct targets of amiRNAs can be accurately predicted using empirically derived determinants of target selection by natural miRNAs (Schwab et al. 2005). This suggests that the specificity of natural plant miRNAs is primarily due to an intrinsic property of the plant RNA silencing machinery, rather than selection against broad-spectrum miRNAs during evolution. It will be interesting to determine which components are responsible for the specificity differences between the RNA silencing machineries of plants and animals.

This work also shows that amiRNAs provide an efficient tool for silencing of endogenous genes. While some of the original publications suggested near 100% efficiency of hairpin constructs generating siRNAs (Chuang and Meyerowitz 2000; Wesley et al. 2001), other publications indicate more variable effects (Kerschen et al. 2004). In any case, the availability of several complementary silencing technologies will be an advantage. In addition, we have not yet explored the simultaneous use of

Table 18: Functionality of artificial microRNAs in different backbones and under different promoters

amiRNA	Promoter	Backbone	T1 analyzed	Phenotypic penetrance
amiR-lfy-1	p35S	MIR172a	>200	~80%
amiR-lfy-1	pLFY	MIR172a	>200	~15%
amiR-lfy-2	p35S	MIR172a	>200	0
amiR-lfy-2	p35S	MIR319a	>200	~20%
amiR-white-1	p35S	MIR172a	>100	>90%
amiR-white-1	p35S	MIR319a	>100	>95%
amiR-white-1	pML1	MIR319a	>100	~50%
amiR-white-1	pAP1	MIR319a	>100	~60%
amiR-white-1	pAlc	MIR319a	6	66%
amiR-white-2	p35S	MIR172a	>100	>80%
amiR-ft-1	p35S	MIR319a	40	>95%
amiR-ft-2	p35S	MIR172a	40	>95%
amiR-ft-2	p35S	MIR319a	40	>95%
amiR-trichome	p35S	MIR319a	>100	~30% (*try cpc* double type)
amiR-trichome	pAlc	MIR319a	6	33%
amiR-mads-1	p35S	MIR172a	>200	>90%
amiR-mads-2	p35S	MIR319a	>200	>95%
amiR-mads-2	pAP1	MIR319a	>200	>70%
amiR-mads-2	pAG	MIR319a	>200	<10%
amiR-yabby-1	p35S	MIR172a	>200	>40%
amiR-yabby-2	p35S	MIR319a	>200	~20%

several amiRNAs against the same target(s), which is commonplace with siRNAs, and which may further increase the efficacy of amiRNAs.

Independently of any differences in silencing efficiency, amiRNAs offer several advantages over conventional RNAi. First, miRNA precursors generally produce only a single effective small RNA of known sequence. In contrast, several siRNAs with undefined 5' and 3' ends are produced as a silencing trigger from hairpin constructs. Therefore, targets of amiRNAs can be more accurately predicted than those of longer hairpin constructs. Second, because miRNA-insensitive variants can be generated that do not differ in the encoded protein sequence of targets (Palatnik

et al. 2003), mutant defects of amiRNA expressing plants can be complemented, which is not easily possible with RNAi plants. Third, because of their exquisite specificity, amiRNAs can possibly be adapted for strand- an allele-specific knockouts. Fourth, as with natural miRNAs, amiRNAs are likely to be particularly useful for targeting groups of closely related genes, including tandemly arrayed genes. About 4,000 genes in *Arabidopsis* are found in tandem arrays (The Arabidopsis Genome Initiative, 2000), and no convenient tool exists for their knockout.

Discussion

Both plant and animal miRNAs function as small single stranded molecules of only ~19-24 nucleotides in length. These are processed from stemloop structures harbored in primary *MIRNA* transcripts (Bartel 2004). Whereas stemloops are limited to ~70-80 nucleotides in animals and ~80-250 nucleotides in plants, complete transcripts are often much longer and contain 5' cap structures and poly-A tails, which are typical for polymerase II transcripts (Lee et al. 2004). Mapping of *MIRNA* transcription starts has shown that plant miRNA stemloops are normally contained within the first exon of primary transcripts and that the length of transcribed sequence between stemloop and transcription start varies greatly between ~30 and more than 600 nucleotides (this work) (Xie et al. 2005). Since stemloop structures are sufficient for *DCL1* dependent small RNA production (this work), the function of *MIRNA* transcripts regions outside the stemloop structure is currently not known. It has been suggest that they serve as a mass of structured RNA to attract DCL1, however, my results from studies of regulatory elements controlling miRNA expression propose that they might also function in refining *MIRNA* expression domains.

Most *Arabidopsis* miRNA families consist of several members (1-14), and the majority of precursors can efficiently produce small RNAs (this work). Only in the case of miR164, developmental defects have been described for single *MIRNA* mutants (Baker et al. 2005; Guo et al. 2005), suggesting that the individual precursors are not fully redundant in their spatial and/or temporal distribution. However, none of the available insertion mutants for *MIRNA*156 and *MIR*172 genes displayed any obvious abnormalities (my observations). MiRNA mediated regulation of target genes is crucial for plant development, since both expression of some miRNA resistant targets as well as knock out mutants of *DCL1* and *AGO1* cause lethality (Golden et al. 2002; Palatnik et al. 2003; Vaucheret et al. 2004).

In order to study the degree of spatial redundancy between individual *MIRNA*156 and 172 family members, I have generated fusions of putative regulatory promoter regions located upstream of the determined transcription starts, to the *GUS* reporter gene. Whereas similar patterns were observed for promoter fragments of two potential *MIR156* precursors, three predicted *MIR172* promoters produced different, yet partially overlapping stainings. However, additional family members need to be studied for conclusive statements on expression redundancy and the observed patterns of reporter gene expression need to be confirmed by additional

methodologies. In the future, these reporters can be used to identify upstream regulatory factors mediating individual *MIRNA* expression.

MiRNA expression has also been studied by northern blotting, in situ hybridization, as well as microRNA microarrays (Aukerman and Sakai 2003; Axtell and Bartel 2005; Chen 2004; Jones-Rhoades and Bartel 2004; Kidner and Martienssen 2004; Mallory et al. 2004a; Palatnik et al. 2003; Reinhart et al. 2002). All three methods reflect abundance of complete miRNA families and cannot distinguish individual miRNA family members, since they rely on hybridization of mature miRNAs to complementary nucleic acids. To test the authenticity of the observed patterns of *GUS* activity for the different putative miRNA promoter, a novel method called stem-loop RT-PCR might be suitable. It amplifies individual miRNA stemloops with high specificity and sensitivity (Chen et al. 2005), however we have not yet tested its performance on our set of miRNAs.

Overexpression of miRNAs from three different miRNA families has produced a range of developmental abnormalities in transgenic *Arabidopsis* plants (refers to this work). These propose distinct functional roles of miRNAs during development, especially since phenotypes were partially overlapping with single or double mutants of target genes. However, these defects were caused by very strong and ectopic expression of miRNAs and therefore not necessarily mimic endogenous miRNA functions. In order to better understand the roles of miRNAs during normal development, it will be of great importance to study loss-of-function mutants. Due to redundancy between individual miRNA family members, multiple mutants have to be generated for most miRNA families. In a mouse system, knock-down of complete miRNA families by injection of synthetic stable antisense RNAs has been successfully used to analyze effects on tissue deplete of individual miRNA families (Krutzfeldt et al. 2005). However no such method has been established in plants.

Despite its artificial nature, overexpression of miRNAs can also identify interesting applications. A substantial increase in side shoot and leaf number caused by overexpression of miR156 (this study) can be observed not only in transgenic *Arabidopsis*, but also in other plant species, such as tobacco (Tsegaye Dabi, Salk Institute) and is most likely also coupled to an increase in biomass, as I have observed in *Arabidopsis*.

The molecular mechanism of miRNA function requires base-pairing between miRNAs and target mRNAs. Since only limited complementarity is sufficient for target recognition in animals, animal miRNAs tend to have hundreds of targets, which are typically inhibited at the translational level (Brennecke et al. 2005; Lewis et al. 2005;

Lim et al. 2005). In contrast, plant miRNAs targets with higher complementarity to their miRNAs have been identified in numerous cases. Since cleavage products consistent with miRNA mediated processing had been isolated for many these targets, the regulatory function of plant miRNAs was proposed to affect target RNA levels. This had been demonstrated for miR319a and its *TCP* target genes, and we observed similar effects for three additional plant miRNAs, miR156, 159 and 164. Our analysis of miRNA overexpressing effects on the complete *Arabidopsis* transcriptome suggests that plant miRNAs generally regulate only few direct target genes (1-11 in *Arabidopsis)* (this work), which contrasts with reports of animal miRNAs (see above).

Taken together, plant miRNAs appear to directly affect expression of distinct key regulatory genes, such that misregulation of these genes can in extreme cases abolish a complete genetic program. In animals on contrary, miRNAs establish an extra level of gene regulation on top of many others and function as a common stimulus to regulate huge numbers of genes simultaneously. Thereby, they establish distinct domains, which lack expression of numerous proteins when compared to the surrounding areas (Stark et al. 2005).

Comparing authentic miRNA targets to transcript that had similar numbers of mismatches, but were unaffected by miRNA overexpression (non-functional sequences), I have established determinants of plant miRNA target selection. These discriminate with only a small degree of remaining ambiguity between targets and non-targets in our small set of miRNAs, and are also applicable to most other plant miRNA-target pairs. They can now be used to easily predict target of new miRNAs with high confidence, especially when miRNAs are not conserved in related species, so that comparative sequence information is not available. Interestingly, the described selectivity of plant miRNAs is not only much higher than that of animal miRNAs (Brennecke et al. 2005; Lewis et al. 2005; Lim et al. 2005), but also than that of synthetic siRNAs when applied to cultured human cells (Jackson et al. 2003), even though siRNAs also function on the transcript level. SiRNA effects on transcripts, which do not contain perfectly complementary target sites are generally termed off-targets. These represent a major problem in therapeutic application of siRNAs, since effects on not perfectly complementary transcripts are not predictable to date (Hannon and Rossi 2004). Target determinants, as we have established for plant miRNAs, would be needed to solve this problem, however it is not known if they exist in such a simple form.

The differences in specificity of target selection between animal and plant small RNA machineries suggest that individual components have acquired different

selectivities. It will be interesting to determine these components and analyze the molecular nature of their target specificity.

Effects of miRNA overexpression can be complicated when additional levels of gene regulation affect target expression at the same time, as the miRNA. Whereas transcript cleavage induces downregulation of most target transcripts in case of miRNAs 156, 159, 164 and 319, steady state levels of most miR172 targets remain constant in miRNA overexpressing plants (this work) (Aukerman and Sakai 2003; Chen 2004), even though transcript cleavage is increased (this work). At the same time, protein levels of the target gene *AP2* are greatly reduced (Aukerman and Sakai 2003; Chen 2004). We have postulated and confirmed that direct or indirect negative feedback regulation of target gene expression is responsible for these effects. This suggests that miR172 targets are simultaneously regulated by transcript cleavage and by translational inhibition. Since protein levels are not easily measurable in large scales, it remains to be determined by individual studies, to which extent other miRNA-target pairs also use both targeting mechanisms. Similarly, further work is required to determine the advantages coupled with simultaneous use of the two targeting mechanisms.

Feedback regulation has also been observed on other components of miRNA biosynthesis and function, as for example *DCL1* and *AGO1* are themselves regulated by miRNAs (Vaucheret et al. 2004; Xie et al. 2003). Together, these findings suggest that miRNA directed regulation of gene expression is crucial for proper plant development and requires highly coordinated regulation itself. Therefore, additional regulatory levels function in target gene and miRNA control, which might compensate for smaller changes in miRNA abundance. These could include the overlapping expression of individual miRNA family members and also feedback regulation of targets, as exemplified by *AP2*.

Plant miRNAs regulate only small numbers of target transcripts (this work), which contrasts with much broader specificities of animal miRNAs and also synthetic siRNAs when applied to animal systems (Brennecke et al. 2005; Jackson et al. 2003; Lewis et al. 2005; Lim et al. 2005). The narrow action spectrum of natural plant miRNAs might reflect only intrinsic properties of the miRNA machinery, or could be caused by selection against miRNAs with broader specificity. To distinguish between the two alternatives, I have generated a series of artificial miRNAs (amiRNAs) designed to target different endogenous mRNAs, and compared their effects to those of natural miRNAs. The accurate prediction of direct amiRNA targets by the parameters of target selection for natural miRNAs (see above) suggests that

extensive base-pairing with targets is required for proper plant miRNA function, so that the plant miRNA machinery itself determines its high selectivity of target selection.

Since amiRNAs of engineered sequences were efficiently produced from endogenous miRNA precursors and regulated target genes with my previously established target determinants, amiRNAs make an effective tool for specific gene silencing in plants.

A related technique called RNAi also uses small RNAs (siRNAs) to interfere with gene expression of target genes, and similar to miRNAs, siRNAs are generated from hairpin precursors. These are artificially engineered from sequences of the target gene and processing by DCL1 produces numerous small RNAs of unknown 5' ends, which can efficiently regulate expression of perfectly complementary targets (Wesley et al. 2001). However, exact sequences of siRNAs cannot be inferred sequences or structures of their precursors, and determinants of their target selection are still unknown. Thus, effects of siRNAs on transcripts other than the perfectly complementary intended targets are not easily predictable. MiRNA precursors, on contrary, normally produce only a single stabilized and functional small RNA, so that expression of amiRNA precursors reproducibly generates easily predictable effects on the plant transcriptome.

AmiRNAs are specifically useful when several related, but not identical target genes need to be regulated. Especially multiple mutants of tandemly arrayed genes are not easily producible with conventional methods, such as crossing individual mutant lines. AmiRNAs are also effective when expressed under tissue-specific or inducible promoters, with limited non-autonomous effects, extending their application to genes, which produce inviable or infertile offspring when mutated. Importantly, amiRNA resistant versions of target genes can be constructed to complement the observed phenotypes of amiRNA expression. AmiRNAs can also be utilized to study silencing of individual splice forms of a gene of interest, when these are sufficiently distant in sequence. Furthermore, strand-specific transcript silencing by amiRNAs should allow discrimination between effects mediated by sense or antisense transcripts. Loss-of-function effects by amiRNA expression can also easily be extended to different genetic backgrounds.

AmiRNAs represent a new methodology for directed gene silencing in plants, that offers several advantages over other techniques. Their usefulness will greatly facilitate future genetic studies in *Arabidopsis* and other plants. Therefore, we have generalized their design principles and integrated all requirements into a web-based

tool (http://wmd.weigelword.org). This service will be available to the plant community and allows the generation of amiRNAs for all the purposes described above.

References

Abbott, A.L., E. Alvarez-Saavedra, E.A. Miska, N.C. Lau, D.P. Bartel, H.R. Horvitz, and V. Ambros. 2005. The let-7 MicroRNA family members mir-48, mir-84, and mir-241 function together to regulate developmental timing in Caenorhabditis elegans. *Dev Cell* **9:** 403-414.

Achard, P., A. Herr, D.C. Baulcombe, and N.P. Harberd. 2004. Modulation of floral development by a gibberellin-regulated microRNA. *Development* **131:** 3357-3365.

Adai, A., C. Johnson, S. Mlotshwa, S. Archer-Evans, V. Manocha, V. Vance, and V. Sundaresan. 2005. Computational prediction of miRNAs in Arabidopsis thaliana. *Genome Res* **15:** 78-91.

Aida, M., T. Ishida, H. Fukaki, H. Fujisawa, and M. Tasaka. 1997. Genes involved in organ separation in Arabidopsis: an analysis of the cup-shaped cotyledon mutant. *Plant Cell* **9:** 841-857.

Alberts, B., A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter. 2002. *Molecular Biology of the Cell (4th Ed.)*.

Allen, E., Z. Xie, A.M. Gustafson, and J.C. Carrington. 2005. microRNA-directed phasing during trans-acting siRNA biogenesis in plants. *Cell* **121:** 207-221.

Allen, E., Z. Xie, A.M. Gustafson, G.H. Sung, J.W. Spatafora, and J.C. Carrington. 2004. Evolution of microRNA genes by inverted duplication of target gene sequences in Arabidopsis thaliana. *Nat Genet* **36:** 1282-1290.

Aravin, A.A., M. Lagos-Quintana, A. Yalcin, M. Zavolan, D. Marks, B. Snyder, T. Gaasterland, J. Meyer, and T. Tuschl. 2003. The small RNA profile during Drosophila melanogaster development. *Dev Cell* **5:** 337-350.

Arazi, T., M. Talmor-Neiman, R. Stav, M. Riese, P. Huijser, and D.C. Baulcombe. 2005. Cloning and characterization of micro-RNAs from moss. *Plant J* **43:** 837-848.

Aukerman, M.J. and H. Sakai. 2003. Regulation of flowering time and floral organ identity by a MicroRNA and its APETALA2-like target genes. *Plant Cell* **15:** 2730-2741.

Axtell, M.J. and D.P. Bartel. 2005. Antiquity of MicroRNAs and Their Targets in Land Plants. *Plant Cell* **17:** 1658-1673.

Bagga, S., J. Bracht, S. Hunter, K. Massirer, J. Holtz, R. Eachus, and A.E. Pasquinelli. 2005. Regulation by let-7 and lin-4 miRNAs results in target mRNA degradation. *Cell* **122:** 553-563.

Baker, C.C., P. Sieber, F. Wellmer, and E.M. Meyerowitz. 2005. The early extra petals1 mutant uncovers a role for microRNA miR164c in regulating petal number in Arabidopsis. *Curr Biol* **15:** 303-315.

Bao, N., K.W. Lye, and M.K. Barton. 2004. MicroRNA binding sites in Arabidopsis class III HD-ZIP mRNAs are required for methylation of the template chromosome. *Dev Cell* **7:** 653-662.

Bartel, D.P. 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116:** 281-297.

Baulcombe, D. 2004. RNA silencing in plants. *Nature* **431:** 356-363.

Baumberger, N. and D.C. Baulcombe. 2005. Arabidopsis ARGONAUTE1 is an RNA Slicer that selectively recruits microRNAs and short interfering RNAs. *Proc Natl Acad Sci U S A* **102:** 11928-11933.

Becker, A. and G. Theissen. 2003. The major clades of MADS-box genes and their role in the development and evolution of flowering plants. *Mol Phylogenet Evol* **29:** 464-489.

Bentwich, I., A. Avniel, Y. Karov, R. Aharonov, S. Gilad, O. Barad, A. Barzilai, P. Einat, U. Einav, E. Meiri, E. Sharon, Y. Spector, and Z. Bentwich. 2005.

Identification of hundreds of conserved and nonconserved human microRNAs. *Nat Genet* **37:** 766-770.

Berezikov, E., V. Guryev, J. van de Belt, E. Wienholds, R.H. Plasterk, and E. Cuppen. 2005. Phylogenetic shadowing and computational identification of human microRNA genes. *Cell* **120:** 21-24.

Billoud, B., R. De Paepe, D. Baulcombe, and M. Boccara. 2005. Identification of new small non-coding RNAs from tobacco and Arabidopsis. *Biochimie*.

Blazquez, M.A., L.N. Soowal, I. Lee, and D. Weigel. 1997. LEAFY expression and flower initiation in Arabidopsis. *Development* **124:** 3835-3844.

Bohnsack, M.T., K. Czaplinski, and D. Gorlich. 2004. Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *Rna* **10:** 185-191.

Bollman, K.M., M.J. Aukerman, M.Y. Park, C. Hunter, T.Z. Berardini, and R.S. Poethig. 2003. HASTY, the Arabidopsis ortholog of exportin 5/MSN5, regulates phase change and morphogenesis. *Development* **130:** 1493-1504.

Bomblies, K., N. Dagenais, and D. Weigel. 1999. Redundant enhancers mediate transcriptional repression of AGAMOUS by APETALA2. *Dev Biol* **216:** 260-264.

Bonnet, E., J. Wuyts, P. Rouze, and Y. Van de Peer. 2004. Detection of 91 potential conserved plant microRNAs in Arabidopsis thaliana and Oryza sativa identifies important target genes. *Proc Natl Acad Sci U S A* **101:** 11511-11516.

Borner, R., G. Kampmann, J. Chandler, R. Gleissner, E. Wisman, K. Apel, and S. Melzer. 2000. A MADS domain gene involved in the transition to flowering in Arabidopsis. *Plant J* **24:** 591-599.

Borsani, O., J. Zhu, P.E. Verslues, R. Sunkar, and J.K. Zhu. 2005. Endogenous siRNAs Derived from a Pair of Natural cis-Antisense Transcripts Regulate Salt Tolerance in Arabidopsis. *Cell* **123:** 1279-1291.

Boutla, A., C. Delidakis, I. Livadaras, M. Tsagris, and M. Tabler. 2001. Short 5'-phosphorylated double-stranded RNAs induce RNA interference in Drosophila. *Curr Biol* **11:** 1776-1780.

Bowman, J.L., J. Alvarez, D. Weigel, E.M. Meyerowitz, and D.R. Smyth. 1993. Control of flower development in *Arabidopsis thaliana* by *APETALA1* and interacting genes. *Development* **119:** 721-743.

Bowman, J.L., D.R. Smyth, and E.M. Meyerowitz. 1989. Genes directing flower development in Arabidopsis. *Plant Cell* **1:** 37-52.

Brennecke, J., A. Stark, R.B. Russell, and S.M. Cohen. 2005. Principles of microRNA-target recognition. *PLoS Biol* **3:** e85.

Busch, M.A., K. Bomblies, and D. Weigel. 1999. Activation of a floral homeotic gene in Arabidopsis. *Science* **285:** 585-587.

Byzova, M., C. Verduyn, D. De Brouwer, and M. De Block. 2004. Transforming petals into sepaloid organs in Arabidopsis and oilseed rape: implementation of the hairpin RNA-mediated gene silencing technology in an organ-specific manner. *Planta* **218:** 379-387.

Cai, X., C.H. Hagedorn, and B.R. Cullen. 2004. Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *Rna* **10:** 1957-1966.

Cardon, G., S. Hohmann, J. Klein, K. Nettesheim, H. Saedler, and P. Huijser. 1999. Molecular characterisation of the Arabidopsis SBP-box genes. *Gene* **237:** 91-104.

Cardon, G.H., S. Hohmann, K. Nettesheim, H. Saedler, and P. Huijser. 1997. Functional analysis of the Arabidopsis thaliana SBP-box gene SPL3: a novel gene involved in the floral transition. *Plant J* **12:** 367-377.

Chen, C., D.A. Ridzon, A.J. Broomer, Z. Zhou, D.H. Lee, J.T. Nguyen, M. Barbisin, N.L. Xu, V.R. Mahuvakar, M.R. Andersen, K.Q. Lao, K.J. Livak, and K.J.

Guegler. 2005. Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Res* **33:** e179.

Chen, X. 2004. A microRNA as a translational repressor of APETALA2 in Arabidopsis flower development. *Science* **303:** 2022-2025.

Chiou, T.J., K. Aung, S.I. Lin, C.C. Wu, S.F. Chiang, and C.L. Su. 2005. Regulation of Phosphate Homeostasis by MicroRNA in Arabidopsis. *Plant Cell*.

Chuang, C.F. and E.M. Meyerowitz. 2000. Specific and heritable genetic interference by double-stranded RNA in Arabidopsis thaliana. *Proc Natl Acad Sci U S A* **97:** 4985-4990.

Dalmay, T., R. Horsefield, T.H. Braunstein, and D.C. Baulcombe. 2001. SDE3 encodes an RNA helicase required for post-transcriptional gene silencing in Arabidopsis. *Embo J* **20:** 2069-2078.

Deyholos, M.K. and L.E. Sieburth. 2000. Separable whorl-specific expression and negative regulation by enhancer elements within the AGAMOUS second intron. *Plant Cell* **12:** 1799-1810.

Dezulian, T., M. Remmert, J.F. Palatnik, D. Weigel, and D.H. Huson. 2005. Identification of plant microRNA homologs. *Bioinformatics*.

Doench, J.G. and P.A. Sharp. 2004. Specificity of microRNA target selection in translational repression. *Genes Dev* **18:** 504-511.

Donis-Keller, H. 1979. Site specific enzymatic cleavage of RNA. *Nucleic Acids Res* **7:** 179-192.

Dunoyer, P., C. Himber, and O. Voinnet. 2005. DICER-LIKE 4 is required for RNA interference and produces the 21-nucleotide small interfering RNA component of the plant cell-to-cell silencing signal. *Nat Genet* **37:** 1356-1360.

Elbashir, S.M., W. Lendeckel, and T. Tuschl. 2001a. RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev* **15:** 188-200.

Elbashir, S.M., J. Martinez, A. Patkaniowska, W. Lendeckel, and T. Tuschl. 2001b. Functional anatomy of siRNAs for mediating efficient RNAi in Drosophila melanogaster embryo lysate. *Embo J* **20:** 6877-6888.

Engstrom, E.M., A. Izhaki, and J.L. Bowman. 2004. Promoter bashing, microRNAs, and Knox genes. New insights, regulators, and targets-of-regulation in the establishment of lateral organ polarity in Arabidopsis. *Plant Physiol* **135:** 685-694.

Enright, A.J., B. John, U. Gaul, T. Tuschl, C. Sander, and D.S. Marks. 2003. MicroRNA targets in Drosophila. *Genome Biol* **5:** R1.

Farh, K.K., A. Grimson, C. Jan, B.P. Lewis, W.K. Johnston, L.P. Lim, C.B. Burge, and D.P. Bartel. 2005. The Widespread Impact of Mammalian MicroRNAs on mRNA Repression and Evolution. *Science*.

Fire, A., S. Xu, M.K. Montgomery, S.A. Kostas, S.E. Driver, and C.C. Mello. 1998. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature* **391:** 806-811.

Floyd, S.K. and J.L. Bowman. 2004. Gene regulation: ancient microRNA target sequences in plants. *Nature* **428:** 485-486.

Forstemann, K., Y. Tomari, T. Du, V.V. Vagin, A.M. Denli, D.P. Bratu, C. Klattenhoff, W.E. Theurkauf, and P.D. Zamore. 2005. Normal microRNA maturation and germ-line stem cell maintenance requires Loquacious, a double-stranded RNA-binding domain protein. *PLoS Biol* **3:** e236.

Frohman, M.A., M.K. Dush, and G.R. Martin. 1988. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc Natl Acad Sci U S A* **85:** 8998-9002.

Fujii, H., T.J. Chiou, S.I. Lin, K. Aung, and J.K. Zhu. 2005. A miRNA involved in phosphate-starvation response in Arabidopsis. *Curr Biol* **15:** 2038-2043.

Gasciolli, V., A.C. Mallory, D.P. Bartel, and H. Vaucheret. 2005. Partially redundant functions of Arabidopsis DICER-like enzymes and a role for DCL4 in producing trans-acting siRNAs. *Curr Biol* **15:** 1494-1500.

Gendrel, A.V. and V. Colot. 2005. Arabidopsis epigenetics: when RNA meets chromatin. *Curr Opin Plant Biol* **8:** 142-147.

Gleave, A.P. 1992. A versatile binary vector system with a T-DNA organisational structure conducive to efficient integration of cloned DNA into the plant genome. *Plant Mol Biol* **20:** 1203-1207.

Golden, T.A., S.E. Schauer, J.D. Lang, S. Pien, A.R. Mushegian, U. Grossniklaus, D.W. Meinke, and A. Ray. 2002. SHORT INTEGUMENTS1/SUSPENSOR1/CARPEL FACTORY, a Dicer homolog, is a maternal effect gene required for embryo development in Arabidopsis. *Plant Physiol* **130:** 808-822.

Grad, Y., J. Aach, G.D. Hayes, B.J. Reinhart, G.M. Church, G. Ruvkun, and J. Kim. 2003. Computational and experimental identification of C. elegans microRNAs. *Mol Cell* **11:** 1253-1263.

Graf, S., D. Strothmann, S. Kurtz, and G. Steger. 2001. HyPaLib: a database of RNAs and RNA structural elements defined by hybrid patterns. *Nucleic Acids Res* **29:** 196-198.

Grewal, S.I. and J.C. Rice. 2004. Regulation of heterochromatin by histone methylation and small RNAs. *Curr Opin Cell Biol* **16:** 230-238.

Grey, F., A. Antoniewicz, E. Allen, J. Saugstad, A. McShea, J.C. Carrington, and J. Nelson. 2005. Identification and characterization of human cytomegalovirus-encoded microRNAs. *J Virol* **79:** 12095-12099.

Grun, D., Y.L. Wang, D. Langenberger, K.C. Gunsalus, and N. Rajewsky. 2005. microRNA target predictions across seven Drosophila species and comparison to mammalian targets. *PLoS Comput Biol* **1:** e13.

Guo, H.S., Q. Xie, J.F. Fei, and N.H. Chua. 2005. MicroRNA Directs mRNA Cleavage of the Transcription Factor NAC1 to Downregulate Auxin Signals for Arabidopsis Lateral Root Development. *Plant Cell* **17:** 1376-1386.

Haley, B. and P.D. Zamore. 2004. Kinetic analysis of the RNAi enzyme complex. *Nat Struct Mol Biol* **11:** 599-606.

Hamilton, A.J. and D.C. Baulcombe. 1999. A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science* **286:** 950-952.

Hannon, G.J. and J.J. Rossi. 2004. Unlocking the potential of the human genome with RNA interference. *Nature* **431:** 371-378.

Hempel, F.D., D. Weigel, M.A. Mandel, G. Ditta, P.C. Zambryski, L.J. Feldman, and M.F. Yanofsky. 1997. Floral determination and expression of floral regulatory genes in Arabidopsis. *Development* **124:** 3845-3853.

Herr, A.J., M.B. Jensen, T. Dalmay, and D.C. Baulcombe. 2005. RNA polymerase IV directs silencing of endogenous DNA. *Science* **308:** 118-120.

Himber, C., P. Dunoyer, G. Moissiard, C. Ritzenthaler, and O. Voinnet. 2003. Transitivity-dependent and -independent cell-to-cell movement of RNA silencing. *Embo J* **22:** 4523-4533.

Hiraguri, A., R. Itoh, N. Kondo, Y. Nomura, D. Aizawa, Y. Murai, H. Koiwa, M. Seki, K. Shinozaki, and T. Fukuhara. 2005. Specific interactions between Dicer-like proteins and HYL1/DRB-family dsRNA-binding proteins in Arabidopsis thaliana. *Plant Mol Biol* **57:** 173-188.

Holen, T., M. Amarzguioui, M.T. Wiiger, E. Babaie, and H. Prydz. 2002. Positional effects of short interfering RNAs targeting the human coagulation trigger Tissue Factor. *Nucleic Acids Res* **30:** 1757-1766.

Horvitz, H.R. and J.E. Sulston. 1980. Isolation and genetic characterization of cell-lineage mutants of the nematode Caenorhabditis elegans. *Genetics* **96:** 435-454.

Humphreys, D.T., B.J. Westman, D.I. Martin, and T. Preiss. 2005. MicroRNAs control translation initiation by inhibiting eukaryotic initiation factor 4E/cap and poly(A) tail function. *Proc Natl Acad Sci U S A* **102:** 16961-16966.

Hunter, C., H. Sun, and R.S. Poethig. 2003. The Arabidopsis heterochronic gene ZIPPY is an ARGONAUTE family member. *Curr Biol* **13:** 1734-1739.

Irish, V.F. and I.M. Sussex. 1990. Function of the apetala-1 gene during Arabidopsis floral development. *Plant Cell* **2:** 741-753.

Irizarry, R.A., B.M. Bolstad, F. Collin, L.M. Cope, B. Hobbs, and T.P. Speed. 2003. Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res.* **31:** e15.

Ishida, T., M. Aida, S. Takada, and M. Tasaka. 2000. Involvement of CUP-SHAPED COTYLEDON genes in gynoecium and ovule development in Arabidopsis thaliana. *Plant Cell Physiol* **41:** 60-67.

Itoh, J.I., A. Hasegawa, H. Kitano, and Y. Nagato. 1998. A recessive heterochronic mutation, plastochron1, shortens the plastochron and elongates the vegetative phase in rice. *Plant Cell* **10:** 1511-1522.

Jackson, A.L., S.R. Bartz, J. Schelter, S.V. Kobayashi, J. Burchard, M. Mao, B. Li, G. Cavet, and P.S. Linsley. 2003. Expression profiling reveals off-target gene regulation by RNAi. *Nat Biotechnol* **21:** 635-637.

Jefferson, R.A. 1989. The GUS reporter gene system. *Nature* **342:** 837-838.

Jing, Q., S. Huang, S. Guth, T. Zarubin, A. Motoyama, J. Chen, F. Di Padova, S.C. Lin, H. Gram, and J. Han. 2005. Involvement of microRNA in AU-rich element-mediated mRNA instability. *Cell* **120:** 623-634.

Johnson, S.M., S.Y. Lin, and F.J. Slack. 2003. The time of appearance of the C. elegans let-7 microRNA is transcriptionally controlled utilizing a temporal regulatory element in its promoter. *Dev Biol* **259:** 364-379.

Johnston, R.J., Jr. and O. Hobert. 2005. A novel C. elegans zinc finger transcription factor, lsy-2, required for the cell type-specific expression of the lsy-6 microRNA. *Development* **132:** 5451-5460.

Jones-Rhoades, M.W. and D.P. Bartel. 2004. Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. *Mol Cell* **14:** 787-799.

Jopling, C.L., M. Yi, A.M. Lancaster, S.M. Lemon, and P. Sarnow. 2005. Modulation of hepatitis C virus RNA abundance by a liver-specific MicroRNA. *Science* **309:** 1577-1581.

Kanno, T., B. Huettel, M.F. Mette, W. Aufsatz, E. Jaligot, L. Daxinger, D.P. Kreil, M. Matzke, and A.J. Matzke. 2005. Atypical RNA polymerase subunits required for RNA-directed DNA methylation. *Nat Genet* **37:** 761-765.

Kanno, T., M.F. Mette, D.P. Kreil, W. Aufsatz, M. Matzke, and A.J. Matzke. 2004. Involvement of putative SNF2 chromatin remodeling protein DRD1 in RNA-directed DNA methylation. *Curr Biol* **14:** 801-805.

Kardailsky, I., V.K. Shukla, J.H. Ahn, N. Dagenais, S.K. Christensen, J.T. Nguyen, J. Chory, M.J. Harrison, and D. Weigel. 1999. Activation tagging of the floral inducer FT. *Science* **286:** 1962-1965.

Kasschau, K.D., Z. Xie, E. Allen, C. Llave, E.J. Chapman, K.A. Krizan, and J.C. Carrington. 2003. P1/HC-Pro, a viral suppressor of RNA silencing, interferes with Arabidopsis development and miRNA unction. *Dev Cell* **4:** 205-217.

Kerschen, A., C.A. Napoli, R.A. Jorgensen, and A.E. Muller. 2004. Effectiveness of RNA interference in transgenic plants. *FEBS Lett* **566:** 223-228.

Khvorova, A., A. Reynolds, and S.D. Jayasena. 2003. Functional siRNAs and miRNAs exhibit strand bias. *Cell* **115:** 209-216.

Kidner, C.A. and R.A. Martienssen. 2004. Spatially restricted microRNA directs leaf polarity through ARGONAUTE1. *Nature* **428:** 81-84.

Kidner, C.A. and R.A. Martienssen. 2005. The developmental role of microRNA in plants. *Curr Opin Plant Biol* **8:** 38-44.

Kim, J., J.H. Jung, J.L. Reyes, Y.S. Kim, S.Y. Kim, K.S. Chung, J.A. Kim, M. Lee, Y. Lee, V. Narry Kim, N.H. Chua, and C.M. Park. 2005. microRNA-directed

cleavage of ATHB15 mRNA regulates vascular development in Arabidopsis inflorescence stems. *Plant J* **42:** 84-94.

Kiriakidou, M., P.T. Nelson, A. Kouranov, P. Fitziev, C. Bouyioukos, Z. Mourelatos, and A. Hatzigeorgiou. 2004. A combined computational-experimental approach predicts human microRNA targets. *Genes Dev* **18:** 1165-1178.

Kirik, V., M. Simon, K. Wester, J. Schiefelbein, and M. Hulskamp. 2004. ENHANCER of TRY and CPC 2 (ETC2) reveals redundancy in the region-specific control of trichome development of Arabidopsis. *Plant Mol Biol* **55:** 389-398.

Klein, J., H. Saedler, and P. Huijser. 1996. A new family of DNA binding proteins includes putative transcriptional regulators of the Antirrhinum majus floral meristem identity gene SQUAMOSA. *Mol Gen Genet* **250:** 7-16.

Koh, S., A.M. Wiles, J.S. Sharp, F.R. Naider, J.M. Becker, and G. Stacey. 2002. An oligopeptide transporter gene family in Arabidopsis. *Plant Physiol* **128:** 21-29.

Koornneef, M., C.J. Hanhart, and J.H. van der Veen. 1991. A genetic and physiological analysis of late flowering mutants in Arabidopsis thaliana. *Mol Gen Genet* **229:** 57-66.

Krutzfeldt, J., N. Rajewsky, R. Braich, K.G. Rajeev, T. Tuschl, M. Manoharan, and M. Stoffel. 2005. Silencing of microRNAs in vivo with 'antagomirs'. *Nature* **438:** 685-689.

Kurihara, Y. and Y. Watanabe. 2004. Arabidopsis micro-RNA biogenesis through Dicer-like 1 protein functions. *Proc Natl Acad Sci U S A* **101:** 12753-12758.

Lagos-Quintana, M., R. Rauhut, J. Meyer, A. Borkhardt, and T. Tuschl. 2003. New microRNAs from mouse and human. *Rna* **9:** 175-179.

Lagos-Quintana, M., R. Rauhut, A. Yalcin, J. Meyer, W. Lendeckel, and T. Tuschl. 2002. Identification of tissue-specific microRNAs from mouse. *Curr Biol* **12:** 735-739.

Lai, E.C. 2002. Micro RNAs are complementary to 3' UTR sequence motifs that mediate negative post-transcriptional regulation. *Nat Genet* **30:** 363-364.

Lai, E.C., P. Tomancak, R.W. Williams, and G.M. Rubin. 2003. Computational identification of Drosophila microRNA genes. *Genome Biol* **4:** R42.

Larkin, R.M., J.M. Alonso, J.R. Ecker, and J. Chory. 2003. GUN4, a regulator of chlorophyll synthesis and intracellular signaling. *Science* **299:** 902-906.

Laufs, P., A. Peaucelle, H. Morin, and J. Traas. 2004. MicroRNA regulation of the CUC genes is required for boundary size control in Arabidopsis meristems. *Development* **131:** 4311-4322.

Lauter, N., A. Kampani, S. Carlson, M. Goebel, and S.P. Moose. 2005. microRNA172 down-regulates glossy15 to promote vegetative phase change in maize. *Proc Natl Acad Sci U S A* **102:** 9412-9417.

Lee, H., S.S. Suh, E. Park, E. Cho, J.H. Ahn, S.G. Kim, J.S. Lee, Y.M. Kwon, and I. Lee. 2000. The AGAMOUS-LIKE 20 MADS domain protein integrates floral inductive pathways in Arabidopsis. *Genes Dev* **14:** 2366-2376.

Lee, R.C. and V. Ambros. 2001. An extensive class of small RNAs in Caenorhabditis elegans. *Science* **294:** 862-864.

Lee, R.C., R.L. Feinbaum, and V. Ambros. 1993. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* **75:** 843-854.

Lee, Y., C. Ahn, J. Han, H. Choi, J. Kim, J. Yim, J. Lee, P. Provost, O. Radmark, S. Kim, and V.N. Kim. 2003. The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425:** 415-419.

Lee, Y., K. Jeon, J.T. Lee, S. Kim, and V.N. Kim. 2002. MicroRNA maturation: stepwise processing and subcellular localization. *Embo J* **21:** 4663-4670.

Lee, Y., M. Kim, J. Han, K.H. Yeom, S. Lee, S.H. Baek, and V.N. Kim. 2004. MicroRNA genes are transcribed by RNA polymerase II. *Embo J* **23:** 4051-4060.

Lemon, W.J., S. Liyanarachchi, and M. You. 2003. A high performance test of differential gene expression for oligonucleotide arrays. *Genome Biol* **4:** R67.

Lewis, B.P., C.B. Burge, and D.P. Bartel. 2005. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **120:** 15-20.

Lewis, B.P., I.H. Shih, M.W. Jones-Rhoades, D.P. Bartel, and C.B. Burge. 2003. Prediction of mammalian microRNA targets. *Cell* **115:** 787-798.

Li, H., L. Xu, H. Wang, Z. Yuan, X. Cao, Z. Yang, D. Zhang, Y. Xu, and H. Huang. 2005a. The Putative RNA-Dependent RNA Polymerase RDR6 Acts Synergistically with ASYMMETRIC LEAVES1 and 2 to Repress BREVIPEDICELLUS and MicroRNA165/166 in Arabidopsis Leaf Development. *Plant Cell* **17:** 2157-2171.

Li, J., Z. Yang, B. Yu, J. Liu, and X. Chen. 2005b. Methylation protects miRNAs and siRNAs from a 3'-end uridylation activity in Arabidopsis. *Curr Biol* **15:** 1501-1507.

Li, Y., W. Li, and Y.X. Jin. 2005c. Computational identification of novel family members of microRNA genes in Arabidopsis thaliana and Oryza sativa. *Acta Biochim Biophys Sin (Shanghai)* **37:** 75-87.

Liljegren, S.J., G.S. Ditta, Y. Eshed, B. Savidge, J.L. Bowman, and M.F. Yanofsky. 2000. SHATTERPROOF MADS-box genes control seed dispersal in Arabidopsis. *Nature* **404:** 766-770.

Lim, L.P., M.E. Glasner, S. Yekta, C.B. Burge, and D.P. Bartel. 2003a. Vertebrate microRNA genes. *Science* **299:** 1540.

Lim, L.P., N.C. Lau, P. Garrett-Engele, A. Grimson, J.M. Schelter, J. Castle, D.P. Bartel, P.S. Linsley, and J.M. Johnson. 2005. Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* **433:** 769-773.

Lim, L.P., N.C. Lau, E.G. Weinstein, A. Abdelhakim, S. Yekta, M.W. Rhoades, C.B. Burge, and D.P. Bartel. 2003b. The microRNAs of Caenorhabditis elegans. *Genes Dev* **17:** 991-1008.

Lingel, A., B. Simon, E. Izaurralde, and M. Sattler. 2003. Structure and nucleic-acid binding of the Drosophila Argonaute 2 PAZ domain. *Nature* **426:** 465-469.

Liu, J., M.A. Carmell, F.V. Rivas, C.G. Marsden, J.M. Thomson, J.J. Song, S.M. Hammond, L. Joshua-Tor, and G.J. Hannon. 2004. Argonaute2 is the catalytic engine of mammalian RNAi. *Science* **305:** 1437-1441.

Liu, J., F.V. Rivas, J. Wohlschlegel, J.R. Yates, R. Parker, and G.J. Hannon. 2005a. A role for the P-body component GW182 in microRNA function. *Nat Cell Biol* **7:** 1161-1166.

Liu, J., M.A. Valencia-Sanchez, G.J. Hannon, and R. Parker. 2005b. MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies. *Nat Cell Biol* **7:** 719-723.

Liu, Q., T.A. Rand, S. Kalidas, F. Du, H.E. Kim, D.P. Smith, and X. Wang. 2003. R2D2, a bridge between the initiation and effector steps of the Drosophila RNAi pathway. *Science* **301:** 1921-1925.

Llave, C., K.D. Kasschau, M.A. Rector, and J.C. Carrington. 2002a. Endogenous and silencing-associated small RNAs in plants. *Plant Cell* **14:** 1605-1619.

Llave, C., Z. Xie, K.D. Kasschau, and J.C. Carrington. 2002b. Cleavage of Scarecrow-like mRNA targets directed by a class of Arabidopsis miRNA. *Science* **297:** 2053-2056.

Lu, C., S.S. Tej, S. Luo, C.D. Haudenschild, B.C. Meyers, and P.J. Green. 2005a. Elucidation of the small RNA component of the transcriptome. *Science* **309:** 1567-1569.

Lu, S., Y.H. Sun, R. Shi, C. Clark, L. Li, and V.L. Chiang. 2005b. Novel and Mechanical Stress-Responsive MicroRNAs in Populus trichocarpa That Are Absent from Arabidopsis. *Plant Cell* **17:** 2186-2203.

Lund, E., S. Guttinger, A. Calado, J.E. Dahlberg, and U. Kutay. 2004. Nuclear export of microRNA precursors. *Science* **303:** 95-98.

Lynn, K., A. Fernandez, M. Aida, J. Sedbrook, M. Tasaka, P. Masson, and M.K. Barton. 1999. The PINHEAD/ZWILLE gene acts pleiotropically in Arabidopsis development and has overlapping functions with the ARGONAUTE1 gene. *Development* **126:** 469-481.

Maine, E.M. 2000. A conserved mechanism for post-transcriptional gene silencing? *Genome Biol* **1:** REVIEWS1018.

Mallory, A.C., D.P. Bartel, and B. Bartel. 2005. MicroRNA-Directed Regulation of Arabidopsis AUXIN RESPONSE FACTOR17 Is Essential for Proper Development and Modulates Expression of Early Auxin Response Genes. *Plant Cell* **17:** 1360-1375.

Mallory, A.C., D.V. Dugas, D.P. Bartel, and B. Bartel. 2004a. MicroRNA regulation of NAC-domain targets is required for proper formation and separation of adjacent embryonic, vegetative, and floral organs. *Curr Biol* **14:** 1035-1046.

Mallory, A.C., B.J. Reinhart, M.W. Jones-Rhoades, G. Tang, P.D. Zamore, M.K. Barton, and D.P. Bartel. 2004b. MicroRNA control of PHABULOSA in leaf development: importance of pairing to the microRNA 5' region. *Embo J* **23:** 3356-3364.

Mandel, M.A., C. Gustafson-Brown, B. Savidge, and M.F. Yanofsky. 1992. Molecular characterization of the Arabidopsis floral homeotic gene APETALA1. *Nature* **360:** 273-277.

Matzke, M., W. Aufsatz, T. Kanno, L. Daxinger, I. Papp, M.F. Mette, and A.J. Matzke. 2004. Genetic analysis of RNA-mediated transcriptional gene silencing. *Biochim Biophys Acta* **1677:** 129-141.

McConnell, J.R., J. Emery, Y. Eshed, N. Bao, J. Bowman, and M.K. Barton. 2001. Role of PHABULOSA and PHAVOLUTA in determining radial patterning in shoots. *Nature* **411:** 709-713.

Mette, M.F., J. van der Winden, M. Matzke, and A.J. Matzke. 2002. Short RNAs can identify new candidate transposable element families in Arabidopsis. *Plant Physiol* **130:** 6-9.

Millar, A.A. and F. Gubler. 2005. The Arabidopsis GAMYB-Like Genes, MYB33 and MYB65, Are MicroRNA-Regulated Genes That Redundantly Facilitate Anther Development. *Plant Cell* **17:** 705-721.

Miyoshi, K., B.O. Ahn, T. Kawakatsu, Y. Ito, J. Itoh, Y. Nagato, and N. Kurata. 2004. PLASTOCHRON1, a timekeeper of leaf initiation in rice, encodes cytochrome P450. *Proc Natl Acad Sci U S A* **101:** 875-880.

Montgomery, M.K., S. Xu, and A. Fire. 1998. RNA as a target of double-stranded RNA-mediated genetic interference in Caenorhabditis elegans. *Proc Natl Acad Sci U S A* **95:** 15502-15507.

Moussian, B., H. Schoof, A. Haecker, G. Jurgens, and T. Laux. 1998. Role of the ZWILLE gene in the regulation of central shoot meristem cell fate during Arabidopsis embryogenesis. *Embo J* **17:** 1799-1809.

Murashige, T. and F. Skoog. 1962. A revised medium for rapid growth and bio-assay with tobacco tissue cultures. *Physiol. Plant.* **15:** 473-497.

Ohler, U., S. Yekta, L.P. Lim, D.P. Bartel, and C.B. Burge. 2004. Patterns of flanking sequence conservation and a characteristic upstream motif for microRNA gene identification. *Rna* **10:** 1309-1322.

Olsen, P.H. and V. Ambros. 1999. The lin-4 regulatory RNA controls developmental timing in Caenorhabditis elegans by blocking LIN-14 protein synthesis after the initiation of translation. *Dev Biol* **216:** 671-680.

Onodera, Y., J.R. Haag, T. Ream, P.C. Nunes, O. Pontes, and C.S. Pikaard. 2005. Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. *Cell* **120:** 613-622.

Paddison, P.J., J.M. Silva, D.S. Conklin, M. Schlabach, M. Li, S. Aruleba, V. Balija, A. O'Shaughnessy, L. Gnoj, K. Scobie, K. Chang, T. Westbrook, M. Cleary, R. Sachidanandam, W.R. McCombie, S.J. Elledge, and G.J. Hannon. 2004. A resource for large-scale RNA-interference-based screens in mammals. *Nature* **428:** 427-431.

Palatnik, J.F., E. Allen, X. Wu, C. Schommer, R. Schwab, J.C. Carrington, and D. Weigel. 2003. Control of leaf morphogenesis by microRNAs. *Nature* **425:** 257-263.

Parizotto, E.A., P. Dunoyer, N. Rahm, C. Himber, and O. Voinnet. 2004. In vivo investigation of the transcription, processing, endonucleolytic activity, and functional relevance of the spatial distribution of a plant miRNA. *Genes Dev* **18:** 2237-2242.

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

Patzel, V., S. Rutz, I. Dietrich, C. Koberle, A. Scheffold, and S.H. Kaufmann. 2005. Design of siRNAs producing unstructured guide-RNAs results in improved RNA interference efficiency. *Nat Biotechnol* **23:** 1440-1444.

Pelaz, S., G.S. Ditta, E. Baumann, E. Wisman, and M.F. Yanofsky. 2000. B and C floral organ identity functions require SEPALLATA MADS-box genes. *Nature* **405:** 200-203.

Pfeffer, S., A. Sewer, M. Lagos-Quintana, R. Sheridan, C. Sander, F.A. Grasser, L.F. van Dyk, C.K. Ho, S. Shuman, M. Chien, J.J. Russo, J. Ju, G. Randall, B.D. Lindenbach, C.M. Rice, V. Simon, D.D. Ho, M. Zavolan, and T. Tuschl. 2005. Identification of microRNAs of the herpesvirus family. *Nat Methods* **2:** 269-276.

Ratcliff, F.G., S.A. MacFarlane, and D.C. Baulcombe. 1999. Gene silencing without DNA. rna-mediated cross-protection between viruses. *Plant Cell* **11:** 1207-1216.

Rehmsmeier, M., P. Steffen, M. Hochsmann, and R. Giegerich. 2004. Fast and effective prediction of microRNA/target duplexes. *Rna* **10:** 1507-1517.

Reinhart, B.J. and D.P. Bartel. 2002. Small RNAs correspond to centromere heterochromatic repeats. *Science* **297:** 1831.

Reinhart, B.J., E.G. Weinstein, M.W. Rhoades, B. Bartel, and D.P. Bartel. 2002. MicroRNAs in plants. *Genes Dev* **16:** 1616-1626.

Reynolds, A., D. Leake, Q. Boese, S. Scaringe, W.S. Marshall, and A. Khvorova. 2004. Rational siRNA design for RNA interference. *Nat Biotechnol* **22:** 326-330.

Rhoades, M.W., B.J. Reinhart, L.P. Lim, C.B. Burge, B. Bartel, and D.P. Bartel. 2002. Prediction of plant microRNA targets. *Cell* **110:** 513-520.

Riechmann, J.L. and E.M. Meyerowitz. 1998. The AP2/EREBP family of plant transcription factors. *Biol Chem* **379:** 633-646.

Roslan, H.A., M.G. Salter, C.D. Wood, M.R. White, K.P. Croft, F. Robson, G. Coupland, J. Doonan, P. Laufs, A.B. Tomsett, and M.X. Caddick. 2001. Characterization of the ethanol-inducible alc gene-expression system in Arabidopsis thaliana. *Plant J* **28:** 225-235.

Rosso, M.G., Y. Li, N. Strizhov, B. Reiss, K. Dekker, and B. Weisshaar. 2003. An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics. *Plant Mol Biol* **53:** 247-259.

Saito, K., A. Ishizuka, H. Siomi, and M.C. Siomi. 2005. Processing of pre-microRNAs by the Dicer-1-Loquacious complex in Drosophila cells. *PLoS Biol* **3:** e235.

Samach, A., H. Onouchi, S.E. Gold, G.S. Ditta, Z. Schwarz-Sommer, M.F. Yanofsky, and G. Coupland. 2000. Distinct roles of CONSTANS target genes in reproductive development of Arabidopsis. *Science* **288:** 1613-1616.

Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. *Molecular Cloning*. Cold Spring Harbor Laboratory, Cold Spring Harbor.

Samols, M.A., J. Hu, R.L. Skalsky, and R. Renne. 2005. Cloning and identification of a microRNA cluster within the latency-associated region of Kaposi's sarcoma-associated herpesvirus. *J Virol* **79:** 9301-9305.

Schellmann, S., A. Schnittger, V. Kirik, T. Wada, K. Okada, A. Beermann, J. Thumfahrt, G. Jurgens, and M. Hulskamp. 2002. TRIPTYCHON and CAPRICE mediate lateral inhibition during trichome and root hair patterning in Arabidopsis. *Embo J* **21:** 5036-5046.

Schmid, M., N.H. Uhlenhaut, F. Godard, M. Demar, R. Bressan, D. Weigel, and J.U. Lohmann. 2003. Dissection of floral induction pathways using global expression analysis. *Development* **130:** 6001-6012.

Schwab, R., J.F. Palatnik, M. Riester, C. Schommer, M. Schmid, and D. Weigel. 2005. Specific effects of microRNAs on the plant transcriptome. *Dev Cell* **8:** 517-527.

Schwarz, D.S., G. Hutvagner, T. Du, Z. Xu, N. Aronin, and P.D. Zamore. 2003. Asymmetry in the assembly of the RNAi enzyme complex. *Cell* **115:** 199-208.

Scortecci, K.C., S.D. Michaels, and R.M. Amasino. 2001. Identification of a MADS-box gene, FLOWERING LOCUS M, that represses flowering. *Plant J* **26:** 229-236.

Sessions, A., D. Weigel, and M.F. Yanofsky. 1999. The Arabidopsis thaliana MERISTEM LAYER 1 promoter specifies epidermal expression in meristems and young primordia. *Plant J* **20:** 259-263.

Smith, T.F. and M.S. Waterman. 1981. Identification of common molecular subsequences. *J. Mol. Biol.* **147:** 195-197.

Song, J.J., J. Liu, N.H. Tolia, J. Schneiderman, S.K. Smith, R.A. Martienssen, G.J. Hannon, and L. Joshua-Tor. 2003. The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes. *Nat Struct Biol* **10:** 1026-1032.

Song, J.J., S.K. Smith, G.J. Hannon, and L. Joshua-Tor. 2004. Crystal structure of Argonaute and its implications for RISC slicer activity. *Science* **305:** 1434-1437.

Souer, E., A. van Houwelingen, D. Kloos, J. Mol, and R. Koes. 1996. The no apical meristem gene of Petunia is required for pattern formation in embryos and flowers and is expressed at meristem and primordia boundaries. *Cell* **85:** 159-170.

Souret, F.F., J.P. Kastenmayer, and P.J. Green. 2004. AtXRN4 degrades mRNA in Arabidopsis and its substrates include selected miRNA targets. *Mol Cell* **15:** 173-183.

Stark, A., J. Brennecke, R.B. Russell, and S.M. Cohen. 2003. Identification of Drosophila MicroRNA targets. *PLoS Biol* **1:** E60.

Stone, J.M., X. Liang, E.R. Nekl, and J.J. Stiers. 2005. Arabidopsis AtSPL14, a plant-specific SBP-domain transcription factor, participates in plant development and sensitivity to fumonisin B1. *Plant J* **41:** 744-754.

Suh, M.R., Y. Lee, J.Y. Kim, S.K. Kim, S.H. Moon, J.Y. Lee, K.Y. Cha, H.M. Chung, H.S. Yoon, S.Y. Moon, V.N. Kim, and K.S. Kim. 2004. Human embryonic stem cells express a unique set of microRNAs. *Dev Biol* **270:** 488-498.

Sunkar, R., T. Girke, P.K. Jain, and J.K. Zhu. 2005. Cloning and Characterization of MicroRNAs from Rice. *Plant Cell* **17:** 1397-1411.

Sunkar, R. and J.K. Zhu. 2004. Novel and stress-regulated microRNAs and other small RNAs from Arabidopsis. *Plant Cell* **16:** 2001-2019.

Timmons, L. and A. Fire. 1998. Specific interference by ingested dsRNA. *Nature* **395:** 854.

Tomari, Y., C. Matranga, B. Haley, N. Martinez, and P.D. Zamore. 2004. A protein sensor for siRNA asymmetry. *Science* **306:** 1377-1380.

Tomari, Y. and P.D. Zamore. 2005. MicroRNA biogenesis: drosha can't cut it without a partner. *Curr Biol* **15:** R61-64.

Unte, U.S., A.M. Sorensen, P. Pesaresi, M. Gandikota, D. Leister, H. Saedler, and P. Huijser. 2003. SPL8, an SBP-box gene that affects pollen sac development in Arabidopsis. *Plant Cell* **15:** 1009-1019.

Vaucheret, H., F. Vazquez, P. Crete, and D.P. Bartel. 2004. The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. *Genes Dev* **18:** 1187-1197.

Vazquez, F., V. Gasciolli, P. Crete, and H. Vaucheret. 2004a. The nuclear dsRNA binding protein HYL1 is required for microRNA accumulation and plant development, but not posttranscriptional transgene silencing. *Curr Biol* **14:** 346-351.

Vazquez, F., H. Vaucheret, R. Rajagopalan, C. Lepers, V. Gasciolli, A.C. Mallory, J.L. Hilbert, D.P. Bartel, and P. Crete. 2004b. Endogenous trans-acting siRNAs regulate the accumulation of Arabidopsis mRNAs. *Mol Cell* **16:** 69-79.

Voinnet, O. 2005. Induction and suppression of RNA silencing: insights from viral infections. *Nat Rev Genet* **6:** 206-220.

Voinnet, O. and D.C. Baulcombe. 1997. Systemic signalling in gene silencing. *Nature* **389:** 553.

Voinnet, O., P. Vain, S. Angell, and D.C. Baulcombe. 1998. Systemic spread of sequence-specific transgene RNA degradation in plants is initiated by localized introduction of ectopic promoterless DNA. *Cell* **95:** 177-187.

Wang, J.F., H. Zhou, Y.Q. Chen, Q.J. Luo, and L.H. Qu. 2004a. Identification of 20 microRNAs from Oryza sativa. *Nucleic Acids Res* **32:** 1688-1695.

Wang, J.W., L.J. Wang, Y.B. Mao, W.J. Cai, H.W. Xue, and X.Y. Chen. 2005. Control of Root Cap Formation by MicroRNA-Targeted Auxin Response Factors in Arabidopsis. *Plant Cell* **17:** 2204-2216.

Wang, X.J., J.L. Reyes, N.H. Chua, and T. Gaasterland. 2004b. Prediction and identification of Arabidopsis thaliana microRNAs and their mRNA targets. *Genome Biol* **5:** R65.

Weigel, D., J. Alvarez, D.R. Smyth, M.F. Yanofsky, and E.M. Meyerowitz. 1992. LEAFY controls floral meristem identity in Arabidopsis. *Cell* **69:** 843-859.

Weigel, D. and J. Glazebrook. 2002. *Arabidopsis: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Weir, I., J. Lu, H. Cook, B. Causier, Z. Schwarz-Sommer, and B. Davies. 2004. CUPULIFORMIS establishes lateral organ boundaries in Antirrhinum. *Development* **131:** 915-922.

Wesley, S.V., C.A. Helliwell, N.A. Smith, M.B. Wang, D.T. Rouse, Q. Liu, P.S. Gooding, S.P. Singh, D. Abbott, P.A. Stoutjesdijk, S.P. Robinson, A.P. Gleave, A.G. Green, and P.M. Waterhouse. 2001. Construct design for efficient, effective and high-throughput gene silencing in plants. *Plant J* **27:** 581-590.

Wightman, B., I. Ha, and G. Ruvkun. 1993. Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. *Cell* **75:** 855-862.

Williams, L., S.P. Grigg, M. Xie, S. Christensen, and J.C. Fletcher. 2005. Regulation of Arabidopsis shoot apical meristem and lateral organ formation by microRNA miR166g and its AtHD-ZIP target genes. *Development* **132:** 3657-3668.

Xie, Q., G. Frugis, D. Colgan, and N.H. Chua. 2000. Arabidopsis NAC1 transduces auxin signal downstream of TIR1 to promote lateral root development. *Genes Dev* **14:** 3024-3036.

Xie, Z., E. Allen, N. Fahlgren, A. Calamar, S.A. Givan, and J.C. Carrington. 2005. Expression of Arabidopsis MIRNA Genes. *Plant Physiol*.

Xie, Z., L.K. Johansen, A.M. Gustafson, K.D. Kasschau, A.D. Lellis, D. Zilberman, S.E. Jacobsen, and J.C. Carrington. 2004. Genetic and functional diversification of small RNA pathways in plants. *PLoS Biol* **2:** E104.

Xie, Z., K.D. Kasschau, and J.C. Carrington. 2003. Negative feedback regulation of Dicer-Like1 in Arabidopsis by microRNA-guided mRNA degradation. *Curr Biol* **13:** 784-789.

Yan, K.S., S. Yan, A. Farooq, A. Han, L. Zeng, and M.M. Zhou. 2003. Structure and conserved RNA binding of the PAZ domain. *Nature* **426:** 468-474.

Yekta, S., I.H. Shih, and D.P. Bartel. 2004. MicroRNA-directed cleavage of HOXB8 mRNA. *Science* **304:** 594-596.

Yi, R., Y. Qin, I.G. Macara, and B.R. Cullen. 2003. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev* **17:** 3011-3016.

Yoo, B.C., F. Kragler, E. Varkonyi-Gasic, V. Haywood, S. Archer-Evans, Y.M. Lee, T.J. Lough, and W.J. Lucas. 2004. A systemic small RNA signaling system in plants. *Plant Cell* **16:** 1979-2000.

Yoshikawa, M., A. Peragine, M.Y. Park, and R.S. Poethig. 2005. A pathway for the biogenesis of trans-acting siRNAs in Arabidopsis. *Genes Dev* **19:** 2164-2175.

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

Zamore, P.D., T. Tuschl, P.A. Sharp, and D.P. Bartel. 2000. RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell* **101:** 25-33.

Zeng, Y., E.J. Wagner, and B.R. Cullen. 2002. Both natural and designed micro RNAs can inhibit the expression of cognate mRNAs when expressed in human cells. *Mol Cell* **9:** 1327-1333.

Zhang, B.H., X.P. Pan, Q.L. Wang, G.P. Cobb, and T.A. Anderson. 2005. Identification and characterization of new plant microRNAs using EST analysis. *Cell Res* **15:** 336-360.

Zhang, H. and B.G. Forde. 1998. An *Arabidopsis* MADS box gene that controls nutrient-induced changes in root architecture. *Science* **279:** 407-409.

Zilberman, D., X. Cao, L.K. Johansen, Z. Xie, J.C. Carrington, and S.E. Jacobsen. 2004. Role of Arabidopsis ARGONAUTE4 in RNA-directed DNA methylation triggered by inverted repeats. *Curr Biol* **14:** 1214-1220.

Zuker, M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* **31:** 3406-3415.

Materials and Methods

General Materials

Standard chemicals and organic solvents were purchased from Sigma (Munich, Germany), Bio-Rad (Munich, Germany), Roth (Karlsruhe, Germany), Merck (Darmstadt, Germany) and Roche (Mannheim, Germany).

Restiction endonucleases and most other DNA modifying enzymes were from Fermentas (St. Leon-Rot, Germany).

DNA Polymerases were purchased from Fermentas (Pfu), Stratagene (USA; Pfu Turbo) and Takara (USA; Extaq).

Oligonucleotides were ordered from MWG (Ebersberg, Germany).

Molecular cloning

Standard buffers, solutions and media were prepared according to Sambrook (Sambrook et al. 1989).

Standard molecular cloning techniques were carried out as described by Sambrook (Sambrook et al. 1989)

Strains used for replication of generated plasmid DNA

Escherichia coli:	DH5α (LifeTechnologies), chemically competent
Agrobacterium tumefaciens:	ASE, electrocompetent. Strain is resistant to chloramphenicol and kanamycin

Extraction of DNA from agarose gels and purification of plasmid DNA for sequencing reactions was carried out using the QIAquick®Gel Extraction kit (Qiagen) or the Wizard®SV Gel and PCR clean-up system (Promega, Germany).

Plasmid minipreps from agrobacteria were done with the Wizard®PlusSV MiniPrep kit (Promega). The lysis time was extended from 5 to 15 minutes.

Final concentration of antibiotics in growht media:

Ampicillin (bacteria)	50µg/ml
Spectinomycin (all bacteria)	100µg/ml
Chloramphenicol (agrobacteria)	25µg/ml
Kanamycin (agrobacteria)	25µg/ml

Plasmids

pBluescript SK (Stratagene), ampicillin resistance

pGEM-Teasy (Promega), ampicillin resistance

pBJ36 (Gleave 1992), ampicillin resistance

pMS37, a derivative of pBJ36 with the CaMV35S promoter inserted blunt ended using the NdeI and SalI sites

pRITA1 (Gleave 1992), ampicillin resistance

pMLBart (Gleave 1992), spectinomycin resistance in bacteria, basta resistance in plants

pART27 (Gleave 1992), spectinomycin resistance in bacteria, kanamycin resistance in plants

A list of all oligonucleotides used for molecular cloning can be found in TableM1

Identifiers of all genes studied are listed in Tables 3, 10 and Supplementary Table1 Sequences of miRNA precursors are available at microrna.sanger.ac.uk/cgi-bin/sequences/browse.pl. Their structures are shown in Supplementary Figure 1 (folding temperature 23°C).

Plant work

Plants were grown at 23°C in continuous light, long-days (16h light, 8h dark), or short days (8h light, 16h dark) at 65% relative humidity.

For growth on soil, seeds were frozen for 2 days at -20°C, stratified at 4°C for 2-3 days in 0.1% agarose and spread on soil. Basta treatment was carried out either by direct application to the water when first soaking the soil (1:2000 of stock solution) or by spraying (1:1000 of stock solution), about 7 days after germination.

For growth on sterile 0.5X MS-agar plates (Murashige and Skoog 1962) with or without 25µg/ml kanamycin, seeds were sterilized with 70% ethanol containing 0.1% triton-X-100 for 15 minutes, shortly washed in 95% ethanol and dried on sterile filter paper in flow hoods. After spreading on agar plates, seeds were stratified at 4°C for 2-3 days.

Transformation of *Arabidopsis* was carried out as described in (Weigel and Glazebrook 2002).

114

All plants but *try cpc* double mutants were of the Columbia-0 (Col-0) ecotype.

mutant alleles: *lfy-12* (Weigel et al. 1992)

gun4-1 (Larkin et al. 2003)

ft-10 is an *ft* null allele with a T-DNA insertion from the GABI-Kat collection, isolation number 290E08 (Rosso et al. 2003)

try cpc (L*er* ecotype) (Schellmann et al. 2002).

Analyses of leaf initiation rates for miR156 overexpressers

Leaves of minimally 5mm length were counted every 2-3 days and the rate of leaf initiation was calculated by dividing the number of leaves counted before first flower buds were visible by the number of days during which the counting took place.

Transient expression of miRNAs and their targets in *Nicotiana benthamiana*

Agrobacterium tumefaciens strains containing transgenes expressing miRNAs and target RNAs were resuspended in infiltration solution at an OD_{600} of ~0.2 and co-infiltrated into *N. benthamiana* leaves using plastic syringes. Co-expression of miRNA and targets used ratios of 10 miRNA to 1 target. Empty plasmids were used as controls. Tissue was harvested after 3 days processed for total RNA extraction (see below).

Infiltration solution: 1mM $MgCl_2$

1mM MES pH5.2

0.15mM Acetosyringone

RNA work

DEPC-treatment

solutions were provided with 0.1% DEPC and stirred overnight before autoclaving.

RNA extractions from plant material

a) RNeasy Mini Kit (Qiagen) according to the manufacturer's protocol

b) Trizol method

0.1g plant material was frozen in liquid nitrogen, ground to fine powder and resuspended in 5ml Trizol. A first centrifugation step at 5000g for 15 minutes pelleted all tissue remnants and the RNA containing supernatant was subjected to double

chloroform extraction (1ml) before precipitating with equal volumes of isopropanol overnight at -20°C. RNA was pelleted by centrifugation at 5000g for 30 minutes and washed once with cold 80% ethanol. The RNA pellet was air-dried briefly and resuspended in DEPC-treated water (50-200µl).

concentrations of RNA were measured photometrically at 260nm.

Small RNA northern blots

RNA was isolated with the Trizol method and 3-20µg were separated after denaturation on a 17% polyacrylamid gel containing 7M urea in 0.5XTBE running buffer at 180V. Staining in ethidium bromide containing running buffer was used to visualize the separated RNA prior to blotting on a Nylon transfer membran (Nytran supercharge, Schleicher&Schüll) using a semi-dry blotting system from biorad (10V for 1hour). The membrane was again stained with ethidium bromide and RNA was closs-linked with a UV stratalinker.

DNA probes of antisense orientation to miRNAs were 5' end-labelled with [γ-^{32}P]-ATP (Amersham, USA) using optkinase (USB, USA) according to the manufacturer's protocol (2-5µl [γ-^{32}P]-ATP). Labelled probes were purified with MicroBio-Spin®6 columns (Bio-Rad) and boiled before use. Membranes were prehybridized for 1h at 38°C in PerfectHyb™Plus hybridization buffer before addition of the denatured probe. Hybridizations were carried out at 38°C overnight and membranes were washed at 50°C using a 2xSSC, 0.2% SDS solution until radioactivity was reduced to less that 300cpm. Exposure to BiomaxMS radiofilms (Kodak) were carried out at -80°C for different times.

Solutions:

20ml polyacrylamide gel:	8.5ml 40% acrylamide:bisacrylamide (37:1)
	8.4g urea
	1ml 10X TBE
	5.6ml DEPC-treated water
	10µl TEMED
	40µl APS

10ml 5X loading dye:	5.75ml glycerol 87%
	0.5ml Tris-Cl pH7.7 to 8 1M
	0.1ml EDTA 0.5M
	3.65ml DEPC-treated water
	bromophenol blue

High molecular weight northern blots

RNA was isolated with the Trizol method and 10-20µg were separated after denaturation on a 1.2% agarose gel containing 1.9% formaldehyde in 1XMOPS running buffer. RNA was visualized by ethidium bromide staining and transferred to a positively charged nylon membrane using 10XSSC transfer solution. The membrane was baked for 3h at 80°C before prehybridizing for 4h at 42°C.

Probes were radiolabelled using $[\alpha^{32}P]$-dCTPs and a commercially available kit (prime-a-gene, Promega). Denatured probes were applied to membranes in hybridization buffer and incubated at 42°C overnight. Washes were carried out at 50°C using a 1XSSC, 0.1% SDS after initial brief rinsing in 2XSSC, 0.2%SDS.

Solutions:

2x sample buffer:	1.6ml formaldehyde (37%) 5ml formamide 0.5ml 20X MOPS 1.6ml dye ethidium bromide ~40µg/ml
dye:	5ml glycerol 0.02ml EDTA 0.5M 3ml water bromophenol blue xylene cyanole FF
prehybridization buffer:	1.1ml water 9ml 20XSSPE 15ml formamide 3ml 50X Denhardt's 0.75ml 20%SDS 0.3ml salmon sperm DNA
hybridization buffer:	1.65ml water 3ml 20XSSPE 5ml formamide 0.25ml 20%SDS 0.1ml salmon sperm DNA

Extration of mRNA from total RNA

RNA was extracted with the Trizol method and ~50µg were used for mRNA isolation using a commercial kit (Oligotex mRNA mini kit, Qiagen).

Generation of 5' RACE libraries from mRNA (Generacer kit, Invitrogen) **and mapping of miRNA cleavage products**

Maximal amounts of mRNA were directly used for ligation of the provided RNA adapter without applying the decapping protocol. This allowed enrichment for mRNA cleavage products, since only those, and not full-length mRNAs contain free 5' phosphates, which are necessary for adapter ligation. The SuperScript™III RT-module was used for reverse transcription using oligo-dT primers. Two nested gene-specific reverse oligonucleotides were used for RACE (Frohman et al. 1988). PCR products were gel-purified, cloned into a plasmid vector, and sequenced.

Quantification of miRNA target cleavage products

This method is based on the template cDNA libraries that were constructed to map target cleavage sites (see above). Chimeric oligonucleotides were disigned with the 5' portion hybridizing to the RNA adapter and the 3' region hybridizing to 4 to 6 nucleotides immediately downstream of the previously mapped cleavage site in the gene of interest. Reverse gene-specific oligonucleotides targeted regions 200 to 250 bases downstream of the ligated adapter. To amplify full-length transcripts, forward oligonucleotides that hybridized upstream of, but close to the cleavage site, were used. Tubulin cDNA was amplified as control. The efficiency of adapter ligation was monitored by amplifying the *SCL6* cleavage product (Llave et al. 2002b), which should be unaffected by the miRNAs investigated in this study.

RT-PCR

RNA was extracted wither with the Trizol or the QIAquick method and 2µg were treated with DNaseI (Fermentas) for 15 minutes. Poly-A RNA was reverse transcribed using oligo-dT primers and a commercially available kit (cloned AMV first strand cDNA synthesis kit, Invitrogen). Second strand synthesis was carried out in the presence of a fluorescent dye (Platimun SYBR green) for quantitative analysis according to the manufacturer's protocol. For semi-quantitative RT-PCR, regular PCR protocols containing a enzyme mix of 1:100 extaq:home-made taq (TÜ-mix) were used.

To standardize for equal template amounts, tubulin was amplified as a reference from all templates. Oligonucleotides for RT-PCR were designed to amplify 150 to 350 nucleotides in the cDNA of interest, if possible spanning an exon-intron boundary.

Microarray analyses

The complete procedure from RNA extraction to hybridization of microarrays is described here: http://www.weigelworld.org/resources/microarray/AtGenExpress/.

I used duplicate Affymetrix ATH1 arrays for all experiments. RNA was extracted from the following tissues and genotypes (all but gun-4 and amiR-white were biological replicates). A wild-type control was added for each condition and tissue.

35S::miR156b	long-day grown inflorescence apices, oldest flowers around stage 7
35S::miR164b	long-day grown inflorescence apices, oldest flowers around stage 7
35S::miR172a	short-day grown vegetative apices (7 days)
35S::miR164b	short-day grown vegetative apices (7 days)
35S::miR159a	long-day grown open flowers
jaw-D	short-day grown leaves (6 weeks)
35S::amiR-lfy-1(MIR172a)	continuous light grown inflorescence apices, oldest flowers around stage 10
35S::amiR-mads-1(MIR172a)	continuous light grown inflorescence apices, oldest flowers around stage 10
35S::amiR-mads-2(MIR319a)	continuous light grown inflorescence apices, oldest flowers around stage 10
35S::amiR-white-1(MIR172a)	continuous light grown seedlings (7days on 0.5XMS agar plates without sucrose)
35S::amiR-white-2(MIR172a)	continuous light grown seedlings (7days on 0.5XMS agar plates without sucrose)

Total RNA processed for each array ranged from 3 µg (for vegetative apices and seedlings) to 5 µg for (amiRNA overexpressing inflorescences) and 7 µg (for flowers and leaves). 12 µg labeled cRNA were used. Normalized expression estimates were obtained using gcRMA (bioconductor.org), a modification of the robust multi-array analysis (RMA) algorithm (Irizarry et al. 2003), and significant changes were calculated using logit-T, which applies statistical testing to probe level data (Lemon et al. 2003).

Microarray data from natural miRNA overexpressers has been deposited with the Gene Expression Omnibus database at the NCBI (www.ncbi.nlm.nih.gov/geo; series accession number GSE2078-GSE2081). AmiRNA microarrays have been submitted to the ArrayExpress database (www.ebi.ac.uk/arrayexpress; accession number E-TABM-63).

Table M1: Oligonucleotide sequences and their applications

template	orientation	sequence 5'->3'	purpose
MIR156a	sense	cag gta cca gat tag gtg cct aca tat ac	amplification of genomic region
MIR156a	antisense	cag gat cca gtt cac caa tat tcc atg tct tc	amplification of genomic region
MIR156a	antisense	gtc caa ctt cag aga tct aca aga tg	5'RACE
MIR156a	antisense	caa gag aag caa gtg caa tg	5'RACE
MIR156b	sense	cag gta cca gta aga cac gtg tag aaa tc	amplification of genomic region
MIR156b	antisense	cag gat cca ctt cag ggt gaa gca cat tag	amplification of genomic region
MIR156b	antisense	gca cac gca aag tta tag ac	5'RACE
MIR156b	antisense	caa ctt tct tct cac aga tct ctc c	5'RACE
MIR156c	sense	gac aaa ttt taa gag aaa cgc ata g	amplification of stemloop
MIR156c	antisense	ggg acc gaa tcg gag ccg gaa tct g	amplification of stemloop
MIR156c	antisense	gca aga gaa gca aat gca tc	5'RACE
MIR156c	antisense	aaa acc cca aga taa cat ttc ata c	5'RACE
MIR156c	sense	cac tcg agg tag act act aga ctc cga g	amplification of promoter
MIR156c	antisense	cag aat tcc tag ggt ttt gac caa atc g	amplification of short promoter
MIR156c	antisense	gga tcc gtt tct atg cgt ttc tct taa aat ttg	amplification of long promoter
MIR156d	sense	cag gta cca ctc gtt acc caa aat gaac	amplification of genomic region
MIR156d	antisense	cag tcg aca ggg agg gag aat tct caa ttt g	amplification of genomic region
MIR156d	antisense	cac gca aaa gca acc ata tac	5'RACE
MIR156d	antisense	gga aat tcc aat tag tcc aga aac cga tg	5'RACE
MIR156d	sense	ctc gag ccg tag tct cgc cag aaa g	amplification of promoter
MIR156d	antisense	gga tcc gtg tcg gct gct tta ctt c	amplification of promoter
MIR156f	sense	cag gta cca gga ttc gtg gta tag tgt tac	amplification of genomic region
MIR156f	antisense	cag gat cca ggc tca tgt tgg aat tcg aat c	amplification of genomic region
MIR164a	sense	cag gta cca gtg gac tga gga gga tta tac	amplification of genomic region
MIR164a	antisense	cag tcg aca ggg ttt agg ttt tct tca ac	amplification of genomic region
MIR164a	antisense	cgt gca aat aag caa atg aga cg	5'RACE

120

MIR164a	antisense	cat gag ggg cgt ttg tag tat gag aac	5'RACE
MIR164b	sense	cag gta cca gac cga aag aat gat gga atg	amplification of genomic region
MIR164b	antisense	cag gat cca cta ata gtg atc taa aag gag	amplification of genomic region
MIR164b	sense	gaa ggt gtg tga tga gca ag	amplification of stemloop
MIR164b	antisense	tca cca agg tgg agt ggt cat g	amplification of stemloop
MIR164b	antisense	gaa cta act cat cca tat cat c	5'RACE
MIR164b	antisense	cct tca tca ttc tct ccg acc ac	5'RACE
MIR172a	sense	aaa aat gga aga cta att tcc gga	amplification of stemloop
MIR172a	antisense	ctg aag aag atc tgg atg gaa tcc	amplification of stemloop short
MIR172a	antisense	agc ttg tgg atc tat taa tgt ctt g	amplification of stemloop long
MIR172a	sense	cat aga gaa ctt tgt gga g	5'RACE
MIR172a	antisense	one of sandip's	5'RACE
MIR172b	sense	cag gta cca cac tgc tca aac tgt tta gg	amplification of genomic region
MIR172b	antisense	cag tcg aca cta atg ctc tcc tgg tat cgt g	amplification of genomic region
MIR172b	sense	tcg gcg gat cca tgg aag aaa gct c	amplification of stemloop
MIR172b	antisense	ttt ctc aag ctt tag gta ttt gta g	amplification of stemloop
MIR172b	sense	ctc ata tac ata tca aaa cc	5'RACE
MIR172b	antisense	one of sandip's	5'RACE
MIR172e	sense	tga ata ggc tag cct ttg gtg gat g	amplification of stemloop
MIR172e	antisense	gac aag agt agc cat gta ttt gct g	amplification of stemloop
MIR172e	sense	gaa ccc ttt tct gcg gat cga g	5'RACE
MIR172e	antisense	ggc tag cct att cat cga gaa cct ag	5'RACE
At4g36920	sense	atg tgg gat cta aac gac gca cca cac c	amplification of csd
At4g36920	antisense	gaa gct tgt aat cca atg ctc cac tca aga agg tct cat gag	amplification of csd
At4g36920	sense	gac aaa tgc tgc agc cag ctc cgg att ctc tcc tca tc	introducing mutations in target site
At4g36920	antisense	gat gag gag aga atc cgg agc tgg ctg cag cat ttg tc	introducing mutations in target site
At4g36920	sense	cac tga cat gga ctg aag gag tag aaa tca g	RACE-RT-PCR
At4g36920	sense	cgg gca gca gca aca ttg gta gcg ga	RT-PCR
At4g36920	antisense	aga gga ggt tgg aag cca ttt gtc tgc	RACE-RT-PCR, RT-PCR
At5g60120	sense	gaa ttc atg ctg gat ctc aat c	amplification of csd
At5g60120	antisense	gga tcc gta atc caa tgc tcc act cta tgg tggt ggt gt gg	amplification of csd
At5g60120	sense	gtt ttc aaa tgc agc cag ctc cgg att ctc act ctc ag	introducing mutations in target site

121

At5g60120	antisense	ctg aga gtg aga atc cgg agc tgg ctg cat ttg aaa ac	introducing mutations in target site
At5g60120	sense	gga aag aac acc aga gaa agg gct tat g	RACE-RT-PCR
At5g60120	antisense	cta tgg tgg tgg ttg tgg gcg gtt cat	RACE-RT-PCR, RT-PCR
At5g67180	sense	ggt acc atg tgg aac ctt aac	amplification of csd
At5g67180	antisense	tct aga gta atc caa tgc tcc act cag gga cga gag	amplification of csd
At5g67180	sense	have been deleted 1788-1792	introducing mutations in target site
At5g67180	antisense		introducing mutations in target site
At2g28550	sense	agg gat gat gag taa ctg ggg atg	RT-PCR
At2g28550	antisense	atg tga aaa atc taa aac cca aat gac	RACE-RT-PCR, RT-PCR
At1g56010	sense	caa ctt tga cca aga acc ctc	RT-PCR
At1g56010	antisense	gag cgt ggc tga ggc tga ac	RT-PCR
At5g07680	sense	ccg cta aga atg aat ggg tga tc	RT-PCR
At5g07680	antisense	gct aag gga agg gtt gct gaa g	RT-PCR
At5g61430	sense	ccg caa aga atg aat ggg tg	RT-PCR
At5g61430	antisense	gtt aag aac agg gct gct gaa gc	RT-PCR
At5g39610	sense	ctc cga cca aga aac cga ag	RT-PCR
At5g39610	antisense	cag aaa ttc aaa acg caa tcc	RT-PCR
At5g39610	sense	att cca aac gca atc caa ttc ttc tgt acc	RACE-PCR
At5g39610	antisense	gag ttt ctt gac cgt caa gca aca gct tca t	RACE-PCR
At5g55930	sense	ggg gta act gtt gat ggc ttc ggt aac ctt at	RT-PCR
At5g55930	antisense	ggt aca cat ctt aaa cct tag tta tat cat ttg	RACE-PCR, RT-PCR
At5g55930	antisense	agc tca aat tac ttt taa ttg gta cac atc	RACE-PCR
At4g00150	sense	ctg aca tgg act gaa gga gta gaa agg c	RACE-RT-PCR
At4g00150	antisense	cta aga ggg ctt ggt tgg agg taa agt tg	RACE-RT-PCR
A	sense	ctgcaaggcgattaagttgggtaac	engineering amiRNA
B	antisense	gcggataacaatttcacacaggaaacag	engineering amiRNA
amiR-lfy-1(172)	I	atgtaacagtgaacgtactgtcgccggcaatcaacgacta	engineering amiRNA
amiR-lfy-1(172)	II	ccggcgacagtacgttcactgttacattttcatagagaac	engineering amiRNA
amiR-lfy-1(172)	III	ctgtcgacagtacgttcactgtttcatctgttgatggacg	engineering amiRNA
amiR-lfy-1(172)	IV	atgaaacagtgaacgtactgtcgacagccaacaacgaccg	engineering amiRNA
amiR-lfy-2(172)	I	atgttacgataaacggttgctcgccggcaatcaacgacta	engineering amiRNA
amiR-lfy-2(172)	II	ccggcgagcaaccgtttatcgtaacattttcatagagaac	engineering amiRNA
amiR-lfy-2(172)	III	ctgtcgagcaaccgtttatcgtatcatctgttgatggacg	engineering amiRNA
amiR-lfy-2(172)	IV	atgatacgataaacggttgctcgacagccaacaacgaccg	engineering amiRNA
amiR-lfy-2(319)	I	gattacgataaacggttgctcgctctctcttttgtattcc	engineering amiRNA
amiR-lfy-2(319)	II	gagcgagcaaccgtttatcgtaatcaaagagaatcaatga	engineering amiRNA
amiR-lfy-2(319)	III	gagcaagcaaccgttaatcgtattcacaggtcgtgatatg	engineering amiRNA

amiR-lfy-2(319)	IV	gaatacgattaacggttgcttgctctacatatatattcct	engineering amiRNA
amiR-white-1(319)	I	gattagtgagaatgttgcgccggtctctcttttgtattcc	engineering amiRNA
amiR-white-1(319)	II	gaccggcgcaacattctcactaatcaaagagaatcaatga	engineering amiRNA
amiR-white-1(319)	III	gaccagcgcaacattgtcactattcacaggtcgtgatatg	engineering amiRNA
amiR-white-1(319)	IV	gaatagtgacaatgttgcgctggtctacatatatattcct	engineering amiRNA
amiR-white-2(172)	I	atgttagtgagaatgttgcgccggcggcaatcaacgacta	engineering amiRNA
amiR-white-2(172)	II	ccgccggcgcaacattctcactaacattttcatagagaac	engineering amiRNA
amiR-white-2(172)	III	ctgacggcgcaacattctcactatcatctgttgatggacg	engineering amiRNA
amiR-white-2(172)	IV	atgatagtgagaatgttgcgccgtcagccaacaacgaccg	engineering amiRNA
amiR-white-2(319)	I	gatttaaccagattttgcgtcgctctctcttttgtattcc	engineering amiRNA
amiR-white-2(319)	II	gagcgacgcaaaatctggttaaatcaaagagaatcaatga	engineering amiRNA
amiR-white-2(319)	III	gagcaacgcaaaatcaggttaattcacaggtcgtgatatg	engineering amiRNA
amiR-white-2(319)	IV	gaattaacctgattttgcgttgctctacatatatattcct	engineering amiRNA
amiR-trichome(319)	I	gatcccattcgatactgctcgcctctctcttttgtattcc	engineering amiRNA
amiR-trichome(319)	II	gaggcgagcagtatcgaatgggatcaaagagaatcaatga	engineering amiRNA
amiR-trichome(319)	III	gaggagagcagtatccaatgggttcacaggtcgtgatatg	engineering amiRNA
amiR-trichome(319)	IV	gaacccattggatactgctctcctctacatatatattcct	engineering amiRNA
amiR-mads-1(172)	I	atgttttggagaaagtgacttgtccggcaatcaacgacta	engineering amiRNA
amiR-mads-1(172)	II	ccggacaagtcactttctccaaaacattttcatagagaac	engineering amiRNA
amiR-mads-1(172)	III	ctgtacaagtcactttctccaaatcatctgttgatggacg	engineering amiRNA
amiR-mads-1(172)	IV	atgatttggagaaagtgacttgtacagccaacaacgaccg	engineering amiRNA
amiR-mads-2(319)	I	gattgttctctatcctcttcagctctctcttttgtattcc	engineering amiRNA
amiR-mads-2(319)	II	gagctgaagaggatagagaacaatcaaagagaatcaatga	engineering amiRNA
amiR-mads-2(319)	III	gagccgaagaggatacagaacattcacaggtcgtgatatg	engineering amiRNA
amiR-mads-2(319)	IV	gaatgttctgtatcctcttcggctctacatatatattcct	engineering amiRNA
amiR-yabby-1(172)	I	atgtactgaaagcttctctgtgggcggcaatcaacgacta	engineering amiRNA
amiR-yabby-1(172)	II	ccgcccacagagaagctttcagtacattttcatagagaac	engineering amiRNA
amiR-yabby-1(172)	III	ctgaccacagagaagctttcagttcatctgttgatggacg	engineering amiRNA
amiR-yabby-1(172)	IV	atgaactgaaagcttctctgtggtcagccaacaacgaccg	engineering amiRNA
amiR-yabby-2(319)	I	gatgtatgctgatgggactctcgtctctcttttgtattcc	engineering amiRNA
amiR-yabby-2(319)	II	gacgagagtcccatcagcatacatcaaagagaatcaatga	engineering amiRNA
amiR-yabby-2(319)	III	gacgcgagtcccatctgcatacttcacaggtcgtgatatg	engineering amiRNA

amiR-yabby-2(319)	IV	gaagtatgcagatgggactcgcgtctacatatatattcct	engineering amiRNA
tubulin	sense	gag cct tac aac gct act ctg tct gtc	RT-PCR
tubulin	antisense	aca cca gac ata gta gca gaa atc aag	RT-PCR
CRC	sense	ctt tgt cgt caa acc tcc tga ga	RT-PCR
CRC	antisense	tca ctt ctt ctc acc gaa tcc caa gcc	RT-PCR
FIL	sense	gat tcc taa agc acc acc cgt ta	RT-PCR
FIL	antisense	cag gag cgt aga acc ctt ctt tc	RT-PCR
At1g23420	sense	aat aaa cca cct gag aag cga ca	RT-PCR
At1g23420	antisense	ctc tct cgg aac cca tta ttg ct	RT-PCR
At4g00180	sense	agc caa tag acc ccc aga gaa g	RT-PCR
At4g00180	antisense	cag ctg aac cgt aaa acc ctt ct	RT-PCR
At1g65480(a)	sense	ggt gga gaa gac ctc agg aa	RT-PCR
At1g65480(a)	antisense	caa ttg tag aaa act gcg gc	RT-PCR
At1g65480(b)	sense	tgg ccg cag ttt tct aca at	RT-PCR
At1g65480(b)	antisense	ctc att ttc ctc ccc ctc tc	RT-PCR
T3		caa tta acc ctc act aaa ggg	sequencing
T7		gta ata cga ctc act ata ggg cg	sequencing
SP6		cta ttt agg tga cac tat aga a	sequencing

Supplementary Material

Supplementary Table 1a: Predicted target genes of artificial microRNAs and expression changes in amiRNA overexpressers compared to wild-type controls by microarray analyses.

Identifier	Common name	Fold reduction (gcRMA analysis)
amiR-mads-1		
At2g45660	MADS-box protein SOC1	2.3
At2g13210	MADS-box protein ANR1	2.2 (A)
At5g65050	MADS-box protein MAF2	2.1
At1g77080	MADS-box protein MAF1	1.4
At4g37940	MADS-box protein	A in wild type
At2g22630	MADS-box protein AGL17	A in wild type
At5g54060	MADS-box protein MAF3	not on ATH1 array
At5g09680	cytochrome b5	not on ATH1 array
amiR-mads-2 (weak lines)		
At1g24260	MADS-box protein AGL9	4.0†
At1g26310	MADS-box protein CAL	1.7
At1g48360	expressed protein	upregulated
At1g51890	leucine-rich repeat protein kinase	absent in analyzed tissue
At1g69120	MADS-box protein APETALA1 (AP1; AGL7)	6.3†
At2g03710	MADS-box protein AGL3	1.5
At2g42830	MADS-box protein AGL5	1.3
At2g45650	MADS-box protein AGL6	3.8† (A)
At3g01990	ACT domain-containing protein (ACR6)	absent in analyzed tissue
At3g02310	MADS-box protein SEPALLATA2 (AGL4; SEP2)	3.7†
At3g45360	hypothetical protein	absent in analyzed tissue
At3g58780	MADS-box protein SHATTERPROOF 1 (AGL1; SHP1)	1.2
At4g11880	MADS-box protein AGL14	absent in analyzed tissue
At4g22950	MADS-box protein AGL19	absent in analyzed tissue
At5g03060	expressed protein	1.6
At5g15800	MADS-box protein SEPALLATA1 (AGL2; SEP1)	2.5†
At5g20240	MADS-box protein PISTILLATA (PI)	1.2
At5g60910	MADS-box protein FRUITFULL (FUL; AGL8)	1.5
At4g04394	hypothetical protein	not on ATH1 array

|	amiR-mads-2 (strong lines)	
At1g24260	MADS-box protein AGL9	10.6†
At1g26310	MADS-box protein CAL	1.2
At1g48360	expressed protein	upregulated
At1g51890	leucine-rich repeat protein kinase	absent in analyzed tissue
At1g69120	MADS-box protein APETALA1 (AP1; AGL7)	6.5†
At2g03710	MADS-box protein AGL3	1.8†
At2g42830	MADS-box protein AGL5	3.9†
At2g45650	MADS-box protein AGL6	5.1† (A)
At3g01990	ACT domain-containing protein (ACR6)	absent in analyzed tissue
At3g02310	MADS-box protein SEPALLATA2 (AGL4; SEP2)	7.3†
At3g45360	hypothetical protein	absent in analyzed tissue
At3g58780	MADS-box protein SHATTERPROOF 1 (AGL1; SHP1)	2.5†
At4g11880	MADS-box protein AGL14	absent in analyzed tissue
At4g22950	MADS-box protein AGL19	absent in analyzed tissue
At5g03060	expressed protein	1.4
At5g15800	MADS-box protein SEPALLATA1 (AGL2; SEP1)	2.9
At5g20240	MADS-box protein PISTILLATA (PI)	2.0†
At5g60910	MADS-box protein FRUITFULL (FUL; AGL8)	1.3
At4g04394	hypothetical protein	not on ATH1 array
	amiR-lfy-1	
At5g61850	LFY	4.6†
	amiR-white-1	
At3g59400	GENOMES UNCOUPLED 4 (GUN4)	5.7†
	amiR-white-2	
At3g59400	GENOMES UNCOUPLED 4 (GUN4)	20.5† (A)

A: absent according to Affymetrix GCOS algorithm

†: significantly changed (logit-T, $p < 0.025$)

Supplementary Table 1b: Predicted target genes of artificial microRNAs and expression changes in amiRNA overexpressers compared to wild-type controls by qualitative RT-PCR.

Identifier	Common name	Reduction
amiR-lfy-2		
At5g61850	LFY	not analyzed
amiR-ft-1		
At1g65480	FLOWERING LOCUS T (FT)	yes
amiR-ft-2		
At1g65480	FLOWERING LOCUS T (FT)	yes
amiR-trichome		
At5g53200	TRIPTYCHON (TRY)	not analyzed
At2g46410	CAPRICE (CPC)	not analyzed
At2g30420	ENHANCER OF TRY AND CPC (ETC2)	not analyzed
amiR-yabby-1		
At1g23420	INNER NO OUTER (INO)	yes
At4g00180	YABBY3 (YAB3)	yes
At3g07060	hypothetical protein	not analyzed
At5g11950	put. lysine decarboxylase	not analyzed
At5g02910	F-box protein	not analyzed
amiR-yabby-2		
At1g69180	CRABS CLAW (CRC)	no

Supplementary Figure 1: Structures of natural MIRNA precursors

Structure were derived from mfold (Zuker 2003). Hybridization energy was 23°C.
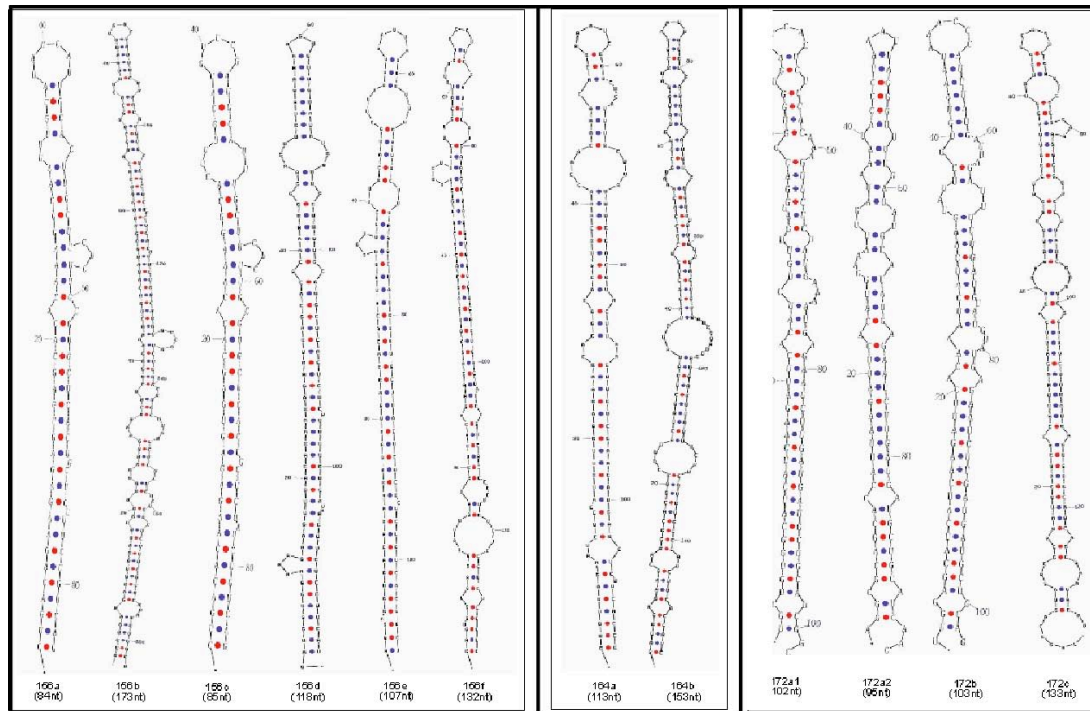
Supplementary Figure 2: Alignments of amiRNAs to target genes.

Energies were derived from mfold (Zuker 2003). Folding energy was calculated at 23°C.

```
5'..CGACACUACGUUCACUGUUA..3'  LFY    At5g61850
    |||||  ||||||||||||||
  3'gcugucaugcaagugacaau 5'   amiR-lfy-1
                                   δG: -31.5 kcal/mol (85.83%)


5'..GGGAGCAUCCGUUUAUCGUAA..3'  LFY    At5g61850
    |  ||||| ||||||||||||||
  3'cgcucguuggcaaauagcauu 5'   amiR-lfy-2
                                   δG: -31.5 kcal/mol (79.15%)


5'..CAACCCUCACCUCCGAGAAUA..3'  FT     At1g65480
    ||||  ||||||||||||||||
  3' aaccaauauuuccuucuccgg 5'  amiR-ft-1
                                   δG: -35.5 kcal/mol (86.59%)


5'..ggcuucuuccuuuauaaccaa..3'  FT     At1g65480
    |||  ||||||||||||||||
  3' ccggagaaggaaauauugguu 5'  amiR-ft-2
                                   δG: -37.0 kcal/mol (94.39%)


5'..CCGCCGCAACAUUCUCACUAA..3'  GUN4   At3g59400
    |||  |||||||||||||||||
  3' ggccgcguuguaagagugauu 5'  amiR-white-1
                                   δG: -34.2 kcal/mol (81.04%)


5'..GCAACGCAAAAUCUGGUUAAA..3'  GUN4   At3g59400
    ||  ||||||||||||||||||
  3' cgcugcguuuuagaccaauuu 5'  amiR-white-2
                                   δG: -32.3 kcal/mol (84.55%)


5'..GGUGAGUAGUAUCGAAUGGGA..3'  CPC    At2g46410
    ||  |||  ||||||||||||||
  3' ccgcucgucauagcuuacccu 5'  amiR-trichome
                                   δG: -40.0 kcal/mol (88.50%)


5'..AGUGAGCAGUAUCGAAUGGGA..3'  TRY    At5g53200
    |  |||||||||||||||||||
  3' ccgcucgucauagcuuacccu 5'  amiR-trichome
                                   δG: -40.6 kcal/mol (89.82%)


5'..AGUGAGUAGCAUCGAAUGGGA..3'  ETC2   At2g30420
    |  |||  ||  ||||||||||
  3' ccgcucgucauagcuuacccu 5'  amiR-trichome
                                   δG: -33.2 kcal/mol (73.45%)


5'..GACAAGUCACUUUCUCCAAAA..3'  AGL17  At2g22630
    ||||||||||||||||||||
  3'cuguucagugaaagagguuuu 5'   amiR-mads-1
                                   δG: -37.5 kcal/mol (100.00%)


5'..GACAAGUCACUUUCUCCAAAC..3'  MAF1   At1g77080
    |||||||||||||||||||
  3'cuguucagugaaagagguuuu 5'   amiR-mads-1
                                   δG: -37.0 kcal/mol (98.67%)


5'..GACAAGUGACUUUCUCCAAAA..3'  SOC1   At2g45660
    |||||||  |||||||||||
  3'cuguucagugaaagagguuuu 5'   amiR-mads-1
                                   δG: -31.8 kcal/mol (84.80%)


5'..GACAAGUCACUUUCUCCAAAC..3'  At4g37940 (MADS-box protein)
    |||||||||||||||||||
  3'cuguucagugaaagagguuuu 5'   amiR-mads-1
                                   δG: -37.0 kcal/mol (98.67%)


5'..GACAAGUCACUUUCUCCAAAC..3'  MAF2   At5g65050
    |||||||||||||||||||
  3'cuguucagugaaagagguuuu 5'   amiR-mads-1
                                   δG: -37.0 kcal/mol (98.67%)
```

```
5'..GACAAGUCACUUUCUCCAAAC..3' MAF3    At5g65060
     ||||||||||||||||||||
3'cuguucagugaaagagguuuu 5'    amiR-mads-1
                                    δG: -37.0 kcal/mol (98.67%)


5'..GACAAGUGACUUUCUCCAAGA..3' ANR1    At2g13210
     |||||||| |||||||||||| |
  3'cuguucagugaaagagguuuu 5'  amiR-mads-1
                                    δG: -31.2 kcal/mol (83.20%)


5'..GACAAGUCUCUUCCUCCAAAG..3' At5g09680     cytochrome b5-domain containing protein
     |||||||| ||| ||||||||
  3'cuguucagugaaagagguuuu 5'  amiR-mads-1
                                    δG: -27.4 kcal/mol (73.07%)


5'..GCUGAAGAGGAUAGAGAACAA..3' AGL3    At2g03710
     |||||||||||||||||||||
  3'cgacuucuccuaucucuuguu 5'  amiR-mads-2
                                    δG: -40.1 kcal/mol (100.00%)


5'..GCUGAAGAGGAUAGAGAACAA..3' AGL2/SEP1    At5g15800
     |||||||||||||||||||||
  3'cgacuucuccuaucucuuguu 5'  amiR-mads-2
                                    δG: -40.1 kcal/mol (100.00%)


5'..GCUGAAGAGGAUAGAGAACAA..3' FUL    At5g60910
     |||||||||||||||||||||
  3'cgacuucuccuaucucuuguu 5'  amiR-mads-2
                                    δG: -40.1 kcal/mol (100.00%)


5'..GAUGAAGAGGAUAGAGAACAA..3' AGL6    At2g45650
     | |||||||||||||||||||
  3'cgacuucuccuaucucuuguu 5'  amiR-mads-2
                                    δG: -35.5 kcal/mol (88.53%)


5'..GCUCAAGAGGAUAGAGAACAA..3'AGL4/SEP2     At3g02310
     ||| ||||||||||||||||
  3'cgacuucuccuaucucuuguu 5'  amiR-mads-2
                                    δG: -34.3 kcal/mol (85.54%)


5'..AUUGAAGAGGAUAGAGAACAA..3' AGL9    At1g24260
     |||||||||||||||||||||
  3'cgacuucuccuaucucuuguu 5'  amiR-mads-2
                                    δG: -35.6 kcal/mol (88.78%)


5'..AUUGAAGAGGAUAGAGAACAA..3' CAL    At1g26310
     |||||||||||||||||||||
  3'cgacuucuccuaucucuuguu 5'  amiR-mads-2
                                    δG: -35.6 kcal/mol (88.78%)


5'..AUUGAAGAGGAUAGAGAACAA..3' AP1    At1g69120
     |||||||||||||||||||||
  3'cgacuucuccuaucucuuguu 5'  amiR-mads-2
                                    δG: -35.6 kcal/mol (88.78%)


5'..GCUGAAGAAGAUGGAGAACAA..3' At5g03060     expressed protein
     ||||||||| ||| ||||||||
  3'cgacuucuccuaucucuuguu 5'  amiR-mads-2
                                    δG: -34.0 kcal/mol (84.79%)


5'..UCAGAAGAGGAUAGAUAACAA..3' At1g51890     LRR protein kinase
     | ||||||||||||| |||||
  3'cgacuucuccuaucucuuguu 5'  amiR-mads-2
                                    δG: -28.8 kcal/mol (71.82%)


5'..GAUAAAGAGGAUAGAGAACAC..3' AGL5    At2g42830
     | | |||||||||||||||||
  3'cgacuucuccuaucucuuguu 5'  amiR-mads-2
                                    δG: -29.1 kcal/mol (72.57%)


5'..GAUAAAGAGGAUAGAGAACAC..3' AGL1/SHP1     At3g58780
     | | |||||||||||||||||
  3'cgacuucuccuaucucuuguu 5'  amiR-mads-2
                                    δG: -29.1 kcal/mol (72.57%)
```

```
5'..GAUGAAGAGGAUAGAGAACGC..3'  AGL14    At4g11880
    |  ||||||||||||||||||
  3'cgacuucuccuaucucuuguu 5'   amiR-mads-2
                                        δG: -34.2 kcal/mol (85.29%)


5'..GAUGAAGAGGAUAGAGAACGC..3'  AGL19    At4g22950
    |  ||||||||||||||||||
  3'cgacuucuccuaucucuuguu 5'   amiR-mads-2
                                        δG: -34.2 kcal/mol (85.29%)


5'..CCAAAAGAGGAUAGAGAGCAA..3'  At3g01990     ACT domain containing prouein
    |   ||||||||||||| |||
  3'cgacuucuccuaucucuuguu 5'   amiR-mads-2
                                        δG: -29.0 kcal/mol (72.32%)


5'..GAUAAAGAGGAUAGAGAACGC..3'  PI       At5g20240
    | |  ||||||||||||||||
  3'cgacuucuccuaucucuuguu 5'   amiR-mads-2
                                        δG: -28.2 kcal/mol (70.32%)


5'..UAUGGAGAAGAUGGAGAACAA..3'  At1g48360     expressed protein
    ||  |||  |||  ||||||||
  3'cgacuucuccuaucucuuguu 5'   amiR-mads-2
                                        δG: -28.2 kcal/mol (70.32%)


5'..ACUGAGGUGGAUGGAGAACAC..3'  At3g45360     hypothetical protein
    ||||  |  ||||  |||||||
  3'cgacuucuccuaucucuuguu 5'   amiR-mads-2
                                        δG: -31.1 kcal/mol (77.56%)


5'..ACUGAGGUGGAUGGAGAACAC..3'  At4g04394     hypothetical protein
    ||||  |  ||||  |||||||
  3'cgacuucuccuaucucuuguu 5'   amiR-mads-2
                                        δG: -31.1 kcal/mol (77.56%)


5'..GCCACAGAGAAGCCUUCAGUG..3'  YAB3     At4g00180
    |||||||||||| ||||||
  3' gggugucucuucgaaagucau 5'  amiR-yabby-1
                                        δG: -35.3 kcal/mol (85.47%)


5'..AUCACAGAGAAGCUAUCAGUG..3'  At3g07060     hypothetical protein
    ||||||||||| |||||
  3' gggugucucuucgaaagucau 5'  amiR-yabby-1
                                        δG: -34.0 kcal/mol (82.32%)


5'..GUCACAGAGAAGUUUUCAGUG..3'  At5g11950     put. lysine decarboxylase
    |||||||||| |||||||
  3' gggugucucuucgaaagucau 5'  amiR-yabby-1
                                        δG: -34.2 kcal/mol (82.81%)


5'..CUCACAAGGAAGCUUUCAGCU..3'  INO      At1g23420
    | |||||   |||||||||||
  3' gggugucucuucgaaagucau 5'  amiR-yabby-1
                                        δG: -30.7 kcal/mol (74.33%)


5'..AACGCAGAGAAGCUUUCACUG..3'  At5g02910     F-box protein
    |  ||||||||||||||| |
  3' gggugucucuucgaaagucau 5'  amiR-yabby-1
                                        δG: -30.4 kcal/mol (73.61%)


5'..AGAGGCUCCCAUCUGCAUACA..3'  CRC      At1g69180
    |||   |||||||  |||||||
  3' gggugucucuucgaaagucau 5'  amiR-yabby-1
                                        δG: -32.0 kcal/mol (73.73%)
```

Lebenslauf

Name:	Rebecca Schwab
Geburtsdatum:	27.01.1979
Geburtsort:	Bietigheim-Bissingen

Schulbildung:

1985 – 1989	Grundschule, Besigheim
1989 – 1998	Christoph-Schrempf-Gymnasium, Besigheim

Studium:

1998 – 2000	Grundstudium der Biologie an der Eberhard Karls Universität Tübingen
2000	Vordiplom in Biologie an der Eberhard Karls Universität Tübingen
2000-2002	Hauptstudium der Biologie an der Eberhard Karls Universität Tübingen
	Fächer: Pflanzenphysiologie
	Genetik
	Biochemie
2001 – 2002	Stipendiat der Studienstiftung des Deutschen Volkes
2001 – 2002	Diplomarbeit im Labor von Prof. Joanne Chory, Salk Institute, La Jolla, USA. Betreuung in Tübingen durch Prof. Wolf Frommer
	Titel: On the function of nuclear speckle formation of Phytochrome B:GFP fusion proteins in *Arabidopsis thaliana*
2002	Diplom in Biologie an der Eberhard Karls Universität Tübingen
10/2002-01/2006	Bearbeitung der vorliegenden Doktorarbeit unter Anleitung von Prof. Detlef Weigel, Max Planck Institut für Entwicklungsbiologie, Tübingen