Adam Mickiewicz University in Poznan
Faculty of Biology
Department of Gene Expression
Uniwersytet im. Adama Mickiewicza w Poznaniu
Wydział Biologii
Zakład Ekspresji Genów

Dawid Bielewicz

Transcriptional and post-transcriptional regulation of microRNAs biogenesis in *Arabidopsis thaliana*

Transkrypcyjna i potranskrypcyjna regulacja biogenezy mikroRNA u *Arabidopsis thaliana*

PhD thesis
Rozprawa doktorska

Poznan, Poland 2015
Poznań, Polska 2015

FUNDING

This work was supported by the Polish National Science Center NCN (UMO-2011/03/N/NZ2/03147), by The Rector's Conference of the Swiss Universities (Sciex Project 11.115) and by KNOW RNA Research Center in Poznan (01/KNOW2/2014).

The PhD fellowship is part of the International PhD Program 'From genome to phenotype: A multidisciplinary approach to functional genomics' (MPD/2010/3) funded by the Foundation for Polish Science (FNP).

SCIENTIFIC COLLABORATION

The experiments were performed at the Department of Gene Expression, Institute of Molecular Biology and Biotechnology, Faculty of Biology, Adam Mickiewicz University, Poznan, Poland as well as the Botanical Institute, University of Basel, Basel, Switzerland.

PUBLICATIONS

The results of the experimental works were described in the following paper:

1. Bielewicz D., Kalak M., Kalna M., Windels D., Barta A., Vazquez F., Szweykowska-Kulińska Z., Jarmołowski A. (2013) Introns of plant pri-miRNA enhance miRNA biogenesis. *EMBO reports* 14(7):622-8.

In addition, the following papers concerning biogenesis and function of miRNAs as well as other small RNAs were published:

1. Kruszka K., Pieczyński M., Windels D., Bielewicz D., Jarmołowski A., Szweykowska-Kulińska Z., Vazquez F. (2012) Role of microRNAs and other sRNAs of plants in their changing environments. *Journal of Plant Physiology* 169(16):164-72.

2. Bielewicz D., Dolata J., Zieleziński A., Alaba S., Szarzyńska B., Szcześniak MW., Jarmołowski A., Szweykowska-Kulińska Z., Karłowski WM. (2011) mirEX: a platform for comparative exploration of plant pri-miRNA expression data. *Nucleic Acids Research* 40(Database issue):D191-7.

ŹRÓDŁA FINANSOWANIA

Niniejsza praca doktorska powstała dzięki dofinansowaniu w ramach grantu uzyskanego z Narodowego Centrum Nauki - NCN (UMO-2011/03/N/NZ2/03147), Konferencji Rektorów Szwajcarskich CRUS (Sciex Project 11.115) oraz z Krajowego Naukowego Ośrodka Wiodącego (KNOW) Poznańskiego Konsorcjum RNA (01/KNOW2/2014).
Projekt doktoratu jest częścią Międzynarodowych Studiów Doktoranckich pt. "From genome to phenotype: A multidisciplinary approach to functional genomics" (MPD/2010/3) uzyskanego z Fundacji na Rzecz Nauki Polskiej (FNP).

WSPÓŁPRACA NAUKOWA

Prace badawcze wykonano w Zakładzie Ekspresji Genów Instytutu Biologii Molekularnej i Biotechnologii Uniwersytetu im. Adama Mickiewicza w Poznaniu oraz w Instytucie Botanicznym Uniwersytetu w Bazylei w Szwajcarii.

PUBLIKACJE

Wyniki badań zostały przedstawione w artykule:

1. Bielewicz D., Kalak M., Kalna M., Windels D., Barta A., Vazquez F., Szweykowska-Kulińska Z., Jarmołowski A. (2013) Introns of plant pri-miRNA enhance miRNA biogenesis. *EMBO reports* 14(7):622-8.

Ponadto następujące prace dotyczące biogenezy i działania miRNA oraz innych małych RNA zostały opublikowane:

1. Kruszka K., Pieczyński M., Windels D., Bielewicz D., Jarmołowski A., Szweykowska-Kulińska Z., Vazquez F. (2012) Role of microRNAs and other sRNAs of plants in their changing environments. *Journal of Plant Physiology* 169(16):164-72.
2. Bielewicz D., Dolata J., Zieleziński A., Alaba S., Szarzyńska B., Szcześniak MW., Jarmołowski A., Szweykowska-Kulińska Z., Karłowski WM. (2011) mirEX: a platform for comparative exploration of plant pri-miRNA expression data. *Nucleic Acids Research* 40(Database issue):D191-7.

SUPERVISORS

PROMOTORZY

Prof. dr hab. Zofia Szweykowska-Kulińska

Department of Gene Expression, Faculty of Biology, Adam Mickiewicz University

Poznan, Poland

Zakład Ekspresji Genów, Wydział Biologii, Uniwersytet im. Adama Mickiewicza

Poznań, Polska

dr Franck Vazquez

Botanical Institute, University of Basel

Basel, Switzerland

Instytut Botaniczny, Uniwersytet w Bazylei

Bazylea, Szwajcaria

REVIEWERS

RECENZENCI

Prof. dr hab. Alicja Józkowicz

Department of Medical Biotechnology, Faculty of Biochemistry, Biophysics and Biotechnology, Jagiellonian University

Krakow, Poland

Zakład Biotechnologii Medycznej, Wydział Biochemii, Biofizyki i Biotechnologii,

Uniwersytet Jagielloński

Kraków, Polska

Prof. dr hab. Andrzej Dziembowski

Institute of Biochemistry and Biophysics, Polish Academy of Science

&

Institute of Genetics & Biotechnology, Warsaw University

Warsaw, Poland

Instytut Biochemii i Biofizyki, Polska Akademia Nauk

oraz

Instytut Genetyki i Biotechnologii, Uniwersytet Warszawski

Warszawa, Polska

ACKNOWLEDGEMENTS

With all of my heart, I would like to thank:

Professor Zofia Szweykowska-Kulińska for giving me the opportunity to be her student and to experience a truly exceptional atmosphere of both hard work and great joy in making scientific discoveries; for her outstanding support, understanding, and great cordiality.

Doctor Franck Vazquez for the warm reception I received in the Botanical Institute laboratory in Basel, the fruitful collaboration, and his valuable comments.

Professor Artur Jarmołowski for the constructive discussions, criticism, and inspiration.

All of my friends, colleagues, co-workers, and professors from Adam Mickiewicz University, especially to the people from the Department of Gene Expression for their invaluable help, constant support, encouragement, inspiration, and great camaraderie.

PODZIĘKOWANIA

Z całego serca pragnę podziękować:

Pani Profesor Zofii Szweykowskiej-Kulińskiej za możliwość bycia Jej studentem i doświadczenia wyjątkowej atmosfery zarówno wytrwałej pracy, jak i wielkiej radości z dokonywania naukowych odkryć; za niezwykłe wsparcie i zrozumienie oraz wielką serdeczność.

Panu Doktorowi Franckowi Vazquezowi za życzliwe przyjęcie w laboratorium Instytutu Botanicznego w Bazylei, owocną współpracę oraz cenne komentarze.

Panu Profesorowi Arturowi Jarmołowskiemu za konstruktywne dyskusje, krytycyzm oraz inspirację.

Wszystkim moim przyjaciołom, znajomym, współpracownikom i profesorom z Uniwersytetu im. Adama Mickiewicza w Poznaniu, w sposób szczególny: wszystkim osobom z Zakładu Ekspresji Genów za ich nieocenioną pomoc, nieustanne wsparcie, dodawanie odwagi, inspirację i wielką serdeczność.

to my wife

mojej żonie

CONTENT

ABBREVIATIONS

A	ADENOSINE
aa	AMINO ACID RESIDUE
ABH1	ABSCISIC ACID HYPERSENSITIVE 1
AGO	ARGONAUTE
APS	AMMONIUM PEROXYDISULFATE
ATP	ADENOSINE 5'-TRIPHOSPHATE
BiFC.	BIMOLECULAR FLUORESCENCE COMPLEMENTATION
CBP	BINDING PROTEIN
CBC	CAP BINDING COMPLEX
CTD	CARBOXY-TERMINAL DOMAIN
DDL	DAWDLE
DCL	DICER LIKE
DRB1	DOUBLE-STRANDED RNA-BINDING PROTEIN 1
dsRBD	DOUBLE-STRANDED RNA-BINDING DOMAIN
dsRNA	DOUBLE-STRANDED RNA
DTT	DITHIOTHREITOL
EDTA	ETHYLENEDIAMINETETRAACETIC ACID
EGTA	ETHYLENE GLYCOL TETRAACETIC ACID
FHA	FORKHEAD-ASSOCIATED DOMAIN
FRET	FLUORESCENCE RESONANCE ENERGY TRANSFER
gDNA	GENOMIC DNA
GFP	GREEN FLUORESCENT PROTEIN
GUS	β-GLUCURONIDASE
HYL1	HYPONASTIC LEAVES 1
miRNA	MICRORNA
PAA	POLYACRYLAMIDE
PAZ	PIWI ARGONAUTE ZWILLE
Pol	POLYMERASE
pri-miRNA	MICRORNA PRIMARY PRECURSOR
RACE	RAPID AMPLIFICATION OF CODING DNA ENDS
ra-siRNA	REPEAT-ASSOCIATED SMALL INTERFERING RNA
RDR6	RNA-DEPENDENT RNA POLYMERASE 6

RT	ROOM TEMPERATURE
RT-PCR	REVERSE TRANSCRIPTION POLYMERASE CHAIN REACTION
SAM	S-ADENOSYLMETHIONINE
SDS	SODIUM DODECYL SULFATE
SE	SERRATE
snRNP	SMALL NUCLEAR RIBONUCLEOPROTEIN PARTICLE
SR	SERINE/ARGININE-RICH
ta-siRNA	*TRANS*-ACTING SMALL INTERFERING RNA
TEMED	N,N,N',N'-TETRAMETHYLETHYLENEDIAMINE
Tris	TRIS[HYDROXYMETHYL]AMINOMETHANE
U snRNA	U-RICH SMALL NUCLEAR RNA
UTR	UNTRANSLATED REGION

ABSTRACT

MicroRNAs have been identified in representatives of almost all groups of eukaryotic organisms and are considered fundamental, sequence-specific regulatory elements in eukaryotic gene expression. Many plant *MIR* genes contain introns. It is known that the presence of introns in protein coding genes leads to alternative splicing in many cases and, as a consequence, a higher diversity of proteins generated from one *locus.* In the case of *MIR* genes, the functional molecule is RNA, 19-24 in length. The presence of introns within *MIR* genes cannot be explained by the ability to increase the diversity of repertoire of microRNAs. This was confirmed by the sequencing of total small RNAs species in *Arabidopsis thaliana*. A question was raised: does the presence of the intron in *MIR* genes have an impact on the maturation of microRNAs? Based on two examples of intron-containing *MIR* genes, *MIR163* and *MIR161*, it was shown that introns are crucial for the accumulation of proper levels of mature microRNA. Removal of the intron led to a drop-off in the level of mature microRNAs. The stimulating effects of the intron mostly reside in the 5′ splice site. Additionally, mutation of the 5' splice site led to a disturbed selection of a polyadenylation site within the primary transcripts of the *MIR* gene. The presence of functional splice sites in the primary transcript of the *MIR163* gene appears mandatory for plant response to biotic stress. The lack of a functional splice site did not stimulate the pathogen-triggered accumulation of miR163 and, as a consequence, led to the improper regulation of its target.

Many proteins are involved in the biogenesis of plant microRNAs. Among them are proteins that recognize double-stranded RNAs and play an important role in the efficient recognition of primary transcripts of microRNA as well as in further steps of the maturation of pri-miRNAs. One of these proteins, the HYL1 (DRB1) protein, interacts with a DCL1 protein, which is the main RNAse enzyme that releases mature microRNAs from their precursors. Recently, it was demonstrated by others that a HYL1 protein has to be dephosphorylated by a CPL1 protein for its optimal activity. On the other hand, CPL1 also dephosphorylates a CTD domain of RNA POL II. These observations suggest that the HYL1 protein is involved in the early steps of biogenesis of microRNA, probable in the initiation of transcription of *MIR* genes. It was shown that expression of the reporter gene that was under control of the *MIR* gene promoter in lack of HYL1 is downregulated. Additionally, higher occupancy of total RNA POL II at several *MIR* gene promoter regions was observed when HYL1 was missing. In the same regions, of *MIR* gene promoters

higher occupancy was also observed in mutants that lacked a functional CPL1 protein. Moreover, we tested an *MIR16*3 gene expression (in the context of HYL1 presence) that was under the control of different RNA POL II promoters. Our results suggest that only when the *MIR163* was under control of its native promoter HYL1 protein influences the level of pri-miRNA163.

STRESZCZENIE

MikroRNA zostały zidentyfikowane u prawie wszystkich przedstawicieli grup organizmów eukariotycznych i są uważane za fundamentalne elementy regulacji ekspresji genów. Wiele roślinnych genów mikroRNA zawiera introny. Wiadomo, że obecność intronów w genach kodujących białka bardzo często prowadzi do alternatywnego splicingu pre-mRNA, a co za tym idzie, do zwiększenia różnorodności powstających białek. Odnośnie genów *MIR* funkcjonalną cząsteczką jest krótki (19-24 nt) RNA. Występowania intronów w takim przypadku nie można wytłumaczyć możliwością zwiększenia repertuaru kodowania różnych mikroRNA. Potwierdzają to wyniki sekwencjonowania całkowitej puli krótkich RNA *Arabidopsis thaliana*. Postawione zostało pytanie: czy obecność intronów w genach i pierwotnych transkryptach tych genów ma wpływ na poziom dojrzałych cząsteczek miRNA? W pracy tej, na podstawie dwóch przykładów genów *MIR* zawierających introny, (*MIR163* i *MIR161*) wykazano, że introny w genach *MIR* są istotne do prawidłowej ekspresji dojrzałych cząsteczek mikroRNA. Usunięcie sekwencji intronu z genu *MIR* prowadzi do obniżenia ekspresji dojrzałej cząsteczki mikroRNA pochodzącej z danego genu. Efekt stymulacji dojrzewania mikroRNA jest związany przede wszystkim z funkcjonalnym miejscem 5' splicingowym. Dodatkowo mutacja miejsca splicingowego 5` prowadzi do zaburzenia w wyborze miejsca poliadenylacji w pierwotnym transkrypcje genu *MIR*. Ponadto obecność miejsc splicingowych w pierwotnym transkrypcie genu *MIR163* okazała się niezbędna w odpowiedzi rośliny na stres biotyczny. Brak miejsc splicingowych podczas ataku patogena nie wywołuje stymulacji ekspresji mikroRNA163, co w konsekwencji prowadzi do nieprawidłowej regulacji genu docelowego tego mikroRNA.

W biogenezę roślinnych mikroRNA zaangażowanych jest wiele białek, między innymi białka wiążące dwuniciowy RNA. Białka te pełnią ważną rolę w procesie prawidłowego i wydajnego rozpoznania pierwotnego transkryptu kodujacego miRNA, a także w dalszych etapach dojrzewania tych cząsteczek. Białko HYL1 (DRB1) oddziałuje z DCL1, główną

RNAzą wycinającą dojrzałe mikroRNA z prekursorów. Wcześniejsze prace wykazały, że defosforylacja białka HYL1 poprzez białko CPL1 jest niezbędna do pełnej aktywności HYL1. Jednocześnie białko CPL1 defosforyluje również domenę CTD polimerazy RNA II. Obserwacje te wskazują na to, że białko HYL1 bierze udział w biogenezie mikroRNA już na najwcześniejszych jej etapach, czyli na etapie inicjacji transkrypcji. W pracy doktorskiej pokazano, że przy braku białka HYL1 ekspresja genu reporterowego, który jest pod kontrolą promotora genu *MIR*, jest obniżona. Dodatkowo zaobserwowaliśmy, że brak białka HYL1 prowadzi do akumulacji polimerazy RNA II w rejonach promotorowych wybranych genów mikroRNA. Polimeraza RNA II również akumuluje się w analizowanych rejonach promotorowych w mutancie z niefunkcjonalnym białkiem CPL1. Ponadto przetestowaliśmy w kontekście obecności białka HYL1 ekspresję genu *MIR*163, który był pod kontrolą różnych promotorów. Uzyskane wyniki sugerują, że tylko w przypadku natywnego promotora *MIR163* białko HYL1 wpływa na poziom pierwotnego transkryptu pri-miRNA163.

INTRODUCTION

1 MicroRNAs

Small RNAs (sRNAs) belong to the regulatory network of gene expression. In plants, these molecules arise through several different metabolic pathways that have several steps in common [Xie *et al.* 2004, Vazquez 2006]. The most abundant class of sRNAs in plants are small interfering RNAs (siRNAs) that are 24 nucleotide (nt) in length [Voinnet 2009]. Small interfering RNAs mainly act at the transcriptional level in plants *via* a process called RNA-directed DNA Methylation (RdDM) [Wierzbicki *et al.* 2008]. The second most abundant class of small RNAs are microRNAs (miRNAs). These molecules are short (19-24 nt in length) endogenous RNAs that control gene expression at the posttranscriptional level. The first miRNA was discovered in a nematode *Caenorhabditis elegans* as a regulator of developmental timing. A *LIN-4* gene that did not encode any protein gives rise to a transcript that is processed into two small RNAs, which are 22 and 61 nucleotides in length, respectively. The 22 nt LIN-4 RNA can be bound to the 3' untranslated region (UTR) of LIN-14 transcripts. *LIN-14* is a protein-coding gene that is also involved in developmental timing in *C. elegans*. Binding between the LIN-4 RNA and LIN-14 transcripts downregulates the level of the LIN-14 protein [Lee *et al.*1993 and 2004]. In 2000, Brenda Reinhart and her colleagues described another microRNA gene in *C. elegans*; *LET-7* [Reinhart *et al.* 2000]. Interestingly, homologs of the *LET-7* gene were also found in other animal species, like humans or fruit flies [Lee & Ambros 2001]. This discovery raised the hypothesis that gene regulation *via* microRNA could be a general mechanism in eukaryotes. Soon thereafter, numerous miRNAs were discovered in various species, such as *Homo sapiens, Mus musculus, Drosophila melanogaster, Arabidopsis thaliana* [Lagos-Quintana *et al.* 2001, Lau *et al.* 2001]. Up till now, more than 300 miRNAs have been identified in *A. thaliana*. However, in *Homo sapiens,* more than 1,000 miRNAs have been found. In *H. sapiens,* microRNAs are predicted to target about 60% of all protein-coding genes [Hafner *et al.* 2012]. Currently, microRNAs have been identified in representatives of all groups of eukaryotic organisms except fungi, and they are considered to be fundamental, sequence-specific regulatory elements in eukaryotic gene expression [Voinnet 2009, Kruszka *et al.* 2012]

1.1 Biogenesis of microRNAs in plants

The overview model of miRNAs biogenesis in *A. thaliana* is shown in Figure 1. The majority of miRNA genes are localized in intergenic regions and are encoded by

independent transcriptional units. In transcription of *MIR* genes, a DNA-dependent RNA polymerase II (POL II), general transcription factors (GTFs), Mediator complex, NEGATIVE ON TATA2 (NOT2), and CELL DIVISION CYCLE5 (CDC5) proteins act as general factors. A multi-subunit complex Mediator was shown to interact with transcriptional activators and facilitate RNA POL II recruitment to *MIR* genes [Kim *et al.* 2011]. NOT2 is a negative transcriptional regulator highly conserved in eukaryotes that also regulates the transcription of protein-coding genes and mRNA decay [Wang *et al.* 2013]. CDC5 physically interacts with RNA POL II and with the promoters of genes that encode miRNAs. This protein positively regulates the transcription of *MIR* genes and affects the occupancy of RNA POL II at *MIR* promoters [Zhang *et al.* 2013]. In the case of particular *MIR* genes, selected transcription factors have been shown to play a regulatory role. It has been shown that transcription factor APETALA2 (AP2) associates with *MIR156* and *MIR172* gene promoters [Yant *et al.* 2010]. Upon copper deficiency, SQUAMOSA PROMOTER BINDING PROTEIN-LIKE7 (SPL7) associates with *MIR398B* and *MIR398C* gene promoters, and upon phosphate deficiency, MYB2 transcription factor binds to the promoter of the *MIR399F* gene [Yamasaki *et al.* 2009, Baek *et al.* 2013].

Primary transcripts of *MIR* genes (pri-miRNAs), like all RNA POL II transcripts in plants, possess both a cap structure and a poly(A) tail. The cap structure of nascent RNA polymerase II transcripts is bound by a CAP-BINDING COMPLEX (CBC), which consists of two proteins: CAP-BINDING PROTEIN20 (CBP20) and CAP BINDING PROTEIN80 (CBP80) [Izaurralde *et al.*1994, Izaurralde *et al.* 1995, Kmieciak *et al.* 2002]. It has been reported that pri-miRNAs generally overaccumulate and the level of mature miRNAs decrease in the *cbp20* and *cbp80* mutants as compared to wild-type plants [Gregory *et al.* 2008, Laubinger *et al.* 2008]. It was also shown that null mutations in the *CBP20* or *CBP80* genes lead to decreased splicing efficiency of cap-proximal intron in a pre-mRNA. Additionally, in the case of a proximal intron, these mutations affect the level of alternatively spliced mRNA isoforms [Kuhn *et al.* 2007, Raczyńska *et al.* 2010]. Both subunits of the CBC are able to interact with another component of microRNA biogenesis machinery called SERRATE (SE). Null mutants of the SE are embryonic lethal, which underlines the importance of this protein in plant development. In a partially loss-of-function *se* mutant, pri-miRNAs generally accumulate, and the level of their mature miRNAs severely decreases [Grigg *et al*. 2005, Lobbes *et al.* 2006, Yang *et al.* 2006a]. SE orthologue ARSENITE RESISTANCE PROTEIN 2 (ARS2) was also reported to play

an essential role in microRNA biogenesis pathways in flies and mammals [Gruber *et al.* 2009, Sabin *et al.* 2009]. Moreover, similar to CBC proteins, the SE protein influences mRNA splicing of a cap-proximal intron and also influences the level of alternative mRNA isoforms [Laubinger *et al.* 2008, Raczyńska *et al.* 2012]. The SE protein, along with two other proteins DICER-LIKE1 (DCL1) and HYPONASTIC LEAVES1 (HYL1), form small nuclear bodies called Dicer-bodies (D-bodies), where it is believed that pri-miRNA processing occurs [Vazquez *et al.* 2004a, Song *et al.* 2007]. DCL1 is an RNase type III endonuclease which catalyzes cleavage reactions of pri-miRNAs through pre-miRNAs till the miRNA/miRNA* duplex is released. In *A. thaliana*, the DCL protein family consists of four proteins, and DCL1 is the key miRNA biogenesis enzyme [Park *et al.* 2002]. The cleavage of pri-miRNAs and pre-miRNAs can be done by DCL1 alone, but the level of efficiency and accuracy of this process requires HYL1 and SE proteins [Dong *et al.* 2008]. Currently, two models of DCL1 action have been proposed. In the first model, dicing of the microRNA primary precursor starts at the base of the stem-loop structure [Kurihara & Watanabe 2004]. In the second model of DCL1 action, the first cut occurs at the site close to the apical loop of a pri-miRNA [Bologna *et al.* 2009]. It looks like some of the pri-miRNAs are processed according to the first model, while the other pri-miRNAs are processed according to the other [Rogers & Chen 2013]. The major difference as compared with animals is the segregated cleavage of miRNA precursors by nuclear and cytoplasmic RNase III enzymes. To liberate pre-miRNA hairpins from pri-miRNA, all animals use the DROSHA enzyme, which interacts with the double-stranded RNA-binding protein DGCR8. Pre-miRNAs are cleaved into miRNA/miRNA* duplexes by the cytoplasmic DICER protein. In humans, DICER is accompanied by the HIV TAR RNA-BINDING (TRBP) protein [Krol *et al.* 2010]. The function of DCL1 may be regulated by a nuclear RNA-binding protein DAWDLE (DDL) that contains a FORKHEAD ASSOCIATED (FHA) domain [Morris *et al.* 2006, Yu *et al.* 2008]. The FHA domain of the DDL protein has a phosphothreonine binding cleft, which can recognize and bind to the phosphothreonine of DCL1 [Machida & Yuan 2013]. Additionally, the phosphorylation status of the HYL1 is also important in microRNA biogenesis. It was shown that the HYL1 protein needs to be dephosphorylated for its optimal activity [Manavella *et al.* 2012]. Hypophosphorylation of the HYL1 is maintained by a C-TERMINAL DOMAIN PHOSPHATASE-LIKE1 (CPL1) protein. CPL1 physically interacts with SE and HYL1 proteins [Manavella *et al.* 2012, Jeong *et al.* 2013]. On the other hand, phosphorylation of TRBP protein by mitogen-activated protein kinase enhanced miRNA production in the

miRNA-generating complex in humans [Paroo *et al*. 2009]. Previously, it was shown that CPL1 is able to dephosphorylate a C-TERMINAL DOMAIN (CTD) of the RNA POL II at serine 5 residue [Koiwa *et al.* 2004]. HYL1 activity can be regulated by two additional proteins – SICKLE (SIC) and MODIFIER OF SNC2 (MOS2) [Zhan *et al.* 2012, Wu *et al.* 2013]. SIC is a proline-rich protein that co-localizes with HYL1 and may act as a partner of HYL1 during the biogenesis of some miRNAs [Zhan *et al.* 2012]. MOS2 is an RNA-binding protein that does not interact with HYL1 but is required for the localization of HYL1 in nuclear dicing bodies. Additionally, MOS2 interacts physically with pri-miRNAs in plants [Wu *et al.* 2013]. It was also demonstrated that another protein – GLYCINE RICH PROTEIN7 (GRP7) – can bind to pri-miRNAs [Koster *et al.* 2014]. GRP7 is an hnRNP-like protein that was previously shown to affect alternative splicing of pre-mRNA. In this respect, it resembles the function of the CBC and SE proteins. Overexpression of GRP7 in *A. thaliana* leads to an altered level of some microRNAs [Koster *et al.* 2014]. Recently, a new player of HYL1 regulation was described. The protein, called so far protease X, can cleave the N-terminal part of the cytoplasmic fraction of HYL1. However, on 1ight exposure, the protein CONSTANS PHOTOMORPHOGENIC 1 (COP1) is translocated to the cytoplasm and suppresses this cleavage. Moreover, in *cop1-4* or *cop1-6* null mutants, the majority of miRNAs as well as HYL1 are reduced in comparison to wild-type plants [Cho *et al.* 2014].

Another component of the DCL1-HYL1-SERRATE complex was described by Guodong Ren and his colleagues in 2012 [Ren *et al.* 2012a]. The TOUGHT (TGH) protein contains a G-patch and a SUPPRESOR OF A WHITE APRICOT (SWAP) domain. It was shown that TGH might regulate the abundance of miRNAs through two activities. First, TGH may modulate the activity of DCL1; and second, the TGH can influence the abundance of pri-miRNAs in HYL1 complexes [Ren *et al.* 2012a]. Recently, Shuxin Zhang and his colleagues described a new component of the pri-miRNA's processing machinery complex, an evolutionary conserved WD-40 protein called PLEIOTROPIC REGULATORY LOCUS1 (PRL1). PRL1 is an RNA-binding protein that is able to associate with pri-miRNAs and may function as a co-factor to DCL1 activity enhancement [Zhang *et al.* 2014]. Additionally, a protein named RECEPTOR FOR ACTIVATED C KINASE (RACK1) interacts with the SE protein [Speth *et al* 2013]. The RACK1 protein has no catalytic activity. However, RACK1 may act as a signal integrator because of its plethora of interaction partners. In a *rack1* mutant, the level of the majority of miRNAs is decreased in comparison to wild-type plants [Speth *et al.* 2013 and 2014]. Released by

DCL1, 21-24 nt double-stranded RNAs undergo methylation by protein HUA1 ENHANCER1 (HEN1). HEN1 is a protein that recognizes 2-nt-long 3' overhangs of the miRNA/miRNA* duplex and catalyzes (in the presence of Mg^{2+}) 2`-O-methylation of the last 3` nucleotides [Boutet *et al.* 2003, Yang *et al.* 2006b, Huang *et al.* 2009]. Methylation of the microRNAs increases their stability. This stabilization is probably a result of an unspecific 3'end polyuridylation inhibition [Ren *et al.* 2012b]. A protein that is able to add uridine moieties in a non-template directed mode is a nucleotide transferase called HEN1 SUPPRESOR1 (HESO1) [Ren *et al.* 2012b]. In its absence, the level of properly-processed active miRNAs increases, and a fraction of U-tailed miRNAs are reduced. Moreover, overexpression of HESO1 in the absence of HEN1 reduces the abundance of microRNAs. This confirms that uridylation can trigger miRNAs degradation [Ren *et al.* 2012b].

The processes described above take place in the nuclear compartment of a cell [Park *et al.* 2005]. After methylation of miRNA/miRNA,* the duplex is transported to a cytoplasm by a HASTY (HST) protein [Bollman *et al*. 2003]. In animals, pre-miRNAs are already transported to a cytoplasm by protein called Exportin-5. Moreover, while plant miRNAs are universally methylated, most products of animal miRNA genes are not. In *A. thaliana* miRNA degradation, the following proteins are involved: SMALL RNA DEGRADING NUCLEASE1, 2, and 3 (SDN1, SDN2 and SDN3). SDNs belong to a family of 3` to 5` exoribonucleases [Ramachandran & Chen 2008].

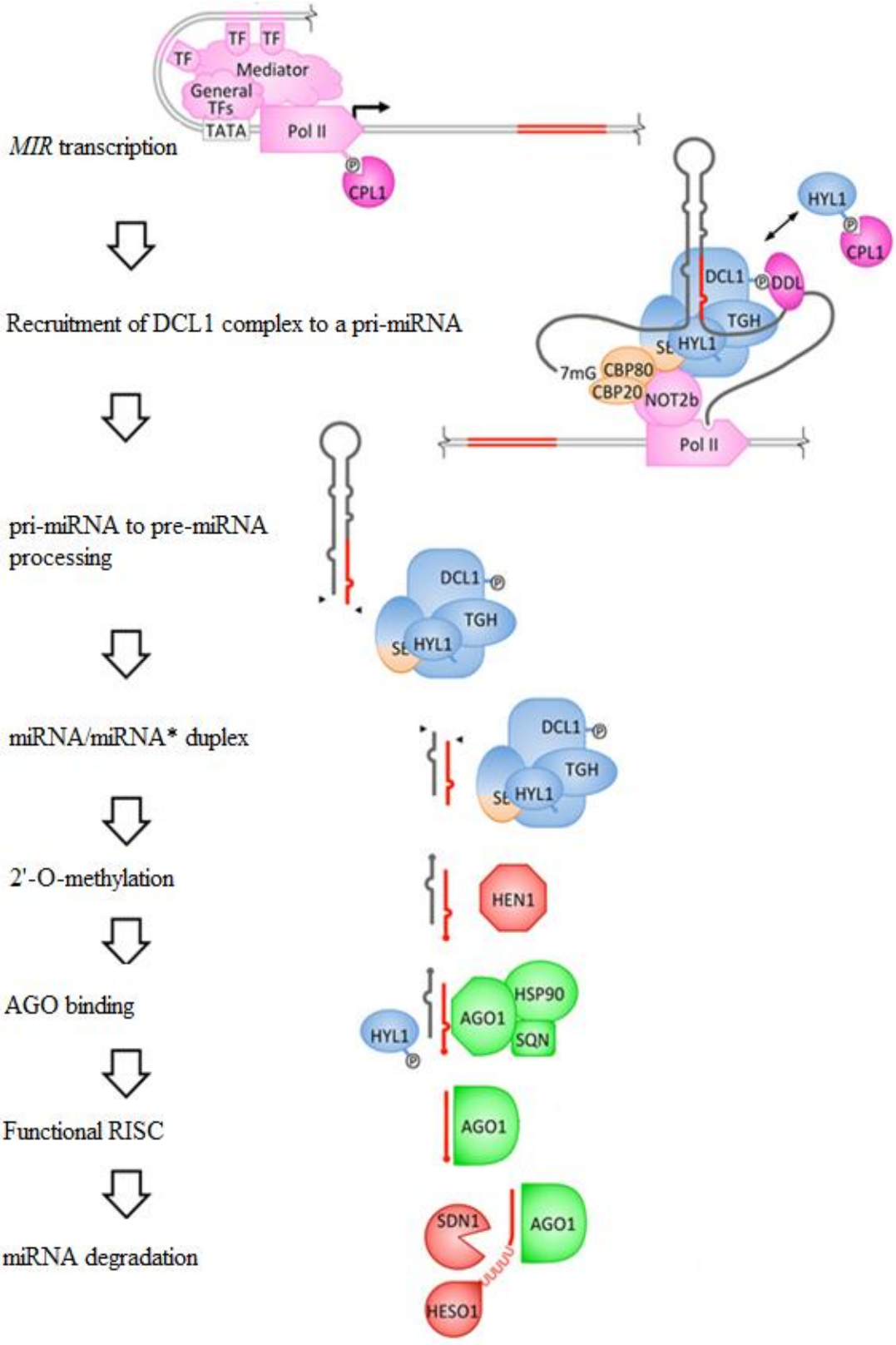


Figure 1 The overview model of miRNA biogenesis in *Arabidopsis* [Rogers and Chen, 2013, modified].

1.2 Mode of action of microRNAs

Mature microRNAs are incorporated into the RNA-INDUCED SILENCING COMPLEX (RISC). At least two proteins are important for the removal of the passenger strand from an miRNA/miRNA* duplex – HEAT SHOCK PROTEIN 90 (HSP90) and SQUINT (SQN). It was shown that HSP90 interacts directly with ARGONAUTE1 (AGO1), which is the most important protein of the RISC. This interaction is driven by ATP-hydrolysis and can trigger conformational changes in AGO1. In addition, the AGO1-associated protein SQN has a similar function as HSP90 in passenger-strand removal [Iki *et al.* 2010, Iki *et al.* 2012]. Additionally, Andrew Eamens and his colleagues showed that, in the absence of the previously mentioned HYL1 protein, incorporation of microRNAs into the RISC complex is deregulated [Eamens *et al.* 2009].

A loaded RISC can downregulate gene expression based on miRNA: mRNA sequence complementarity by two mechanisms: cleavage of the target mRNA or its translational repression. In the first mechanism, one domain of AGO1 forms an RNasH-like fold. This domain has slicer endonuclease activity and is able to cleave RNA targets that are complementary to the miRNA loaded onto RISC [Parker *et al.* 2004, Rivas *et al.* 2005]. In plants, high complementarity between miRNA:mRNA is required for effective target slicing by the AGO1 protein [Mallory *et al.* 2004]. Through cleavage of target RNA, some microRNAs can be involved in the biogenesis of another class of small RNAs called *trans*-acting interfering RNAs (ta-siRNAs) [Allen *et al.* 2005, Gasciolli *et al.* 2005]. TAS RNA, which is generated by RNA POLII, is cleaved in an miRNA-guided manner by an AGO4protein. Next, the cleavage product is converted into double-stranded RNA by an RNA-DEPENDENT RNA POLYMERASE6 (RDR6). Double-stranded RNA is finally processed by the DCL4 into 21-nucleotide-in-size ta-siRNAs [Vazquez *et al.* 2004b, Xie *et al.* 2005a].

In the second mechanism of miRNA action, a decreased level of proteins is maintained through translation inhibition. Initially, it was shown that overexpression of miR172 leads to a decrease of the APETALA2 protein without affecting the abundance of AP2 transcripts [Aukerman & Sakai, 2003]. In fact, in *Arabidopsis* miR172-guided cleavage fragments of AP2 transcripts can be found, which suggests that miR172 is able to function *via* translational inhibition and *via* transcript cleavage [Aukerman & Sakai 2003]. In plants, the detailed mechanism of translation inhibition is still unknown. Recently, it was shown that AGO1 is partially co-localized with endoplasmic reticulum (ER). AGO1 interacts with the ER integral membrane protein ALTERED MERISTEM PROGRAM1 (AMP1) [Li *et*

al. 2013]. It was demonstrated that microRNAs inhibit protein synthesis of their target genes in an AMP1-dependent manner [Li *et al.* 2013]. Additionally, it was shown that cleavage or translational inhibition of microRNA target genes is determined by the DCL1 partnering protein. DRB1 is required for miRNA-guided cleavage, and DRB2 determines miRNA-guided translational inhibition [Reis *et al*. 2015].

2 *Arabidopsis thaliana* microRNA coding units

Generally, *MIR* genes can be classified as independent transcriptional units (intergenic) or as microRNA residing within introns of host-genes (intronic). MiRNAs encoded within the intergenic *loci* are transcribed from their own promoters. Intronic miRNAs are located within introns of host-genes and are thought to be derived from introns of their host genes. In contrast to metazoan, where microRNAs are mainly found within introns or exons of previously characterized genes, the majority of plant microRNAs are encoded by *MIR* genes representing independent transcriptional units [Rodriguez *et al.* 2004; Xie *et al.* 2005b, Szarzyńska *et al*. 2009, Szweykowska-Kulińska *et al.* 2013].

2.1 Genes hosting miRNAs within their introns

The gene organization of eleven intronic (or intragenic) microRNAs was reviewed by John Brown in 2008 [Brown *et al.* 2008]. Intronic microRNAs can be classified into two categories: microRNAs encoded in introns located within the coding part of a host gene or microRNAs encoded in introns located in the untranslated region (UTR) of a host gene (Table 1) [Brown *et al.* 2008].

	name of gene	**microRNA**	**schematic organization**
1st intron	At1g18880	miR837	
	At1g20860	miR416	
	At1g77230	miR402	
2nd intron	At5g08185	miR162a	

	At4g13495	miR850	
3rd intron	At3g23325	miR853	
5th intron	At1g14500	miR852	
14th intron	At1g01040	miR838	
	At2g25170	miR862	
5`UTR	At2g23348	miR844	
3`UTR	At5g13890	miR848	

Table 1 *A. thaliana* microRNA localization within introns of a given host gene. Black boxes represent UTRs, green boxes represent exons, black lines represent introns and blue boxes represent miRNAs [Brown *et al.* 2008, modified].

Plant genes hosting miRNAs within their introns show great diversity of exon/intron organization. Among them, there are single intron-containing genes (At1g18880, At1g20860, At1g77230) as well as multi-intron genes (At1g01040 and At2g25170, with 20 and 30 introns, respectively). The length of an intron where a given miRNA is embedded varies greatly, from several hundred (259 bp, in the case of At5g08185) to over three thousand (2,620 bp, in the case of At4g13495). It seems like there is no rule regarding the localization of miRNA-containing introns in the overall gene structure, as there are examples of miRNA sequences identified within the introns of UTRs and within the introns located in a coding region. Out of all of the microRNAs presented in Table 1, five are encoded within the introns of genes with no annotated function (miR162a, miR844, miR848, miR850, and miR852). In the case of six others (miR402, miR416, miR837, miR838, miR853, and miR862), miRNA sequences overlap with the introns of protein-coding genes.

2.2 Intergenic *MIR* genes

The structure of several *Arabidopsis MIR* genes that are located in intergenic regions was previously described by Bogna Szarzyńska and her colleagues [Szarzyńska *et al.* 2009]. They established the intron/exon organization of twenty microRNA genes. The advantage of Szarzyńska`s approach in determining the structure of *MIR* genes was to use cDNA from the *hyl1-2* mutant where primary transcripts of microRNA genes accumulate. More than half of the analyzed *MIR* genes contained at least one intron (see Table 2), and many transcripts of these *MIR* genes have multiple poly(A) signals. Moreover, it has been previously reported that, in the case of *MIR163, MIR164, MIR164b, MIR164c and MIR171a* genes, transcripts contain alternative polyadenylation sites [Kurihara & Watanabe 2004; Nicovics *et al.* 2006; Song *et al.* 2007].

microRNA	length of *MIR* gene (bp)	number of introns
miR156a	3108	3
miR156c	2580	3
miR157c	997	1
miR158a	535	1
miR159a	808	0
miR160a	2034	1
miR160b	378	0
miR161	699	1
miR164c	832	0
miR166a	1113	1
miR166b	1121	2
miR167a	602	0
miR169f	735	0
miR171b	775	1
miR171c	861	1
miR172a	2097	2
miR172b	1417	3
miR172e	817	0
miR319b	873	0
miR393a	546	0

Table 2 *A. thaliana MIR* genes with fully described structure [Szarzynska *et al.* 2009, modified]

3 Splicing machinery and splicing

In a chemical sense, splicing of introns from mRNA precursors is a simple double transesterification reaction that can be divided into two steps, both involving a single transesterification reaction. The first step is the cleavage of the 5′ splice site. This occurs by a transesterification reaction promoted by the 2’ hydroxyl group of a special adenosine nucleotide located within the intron sequence called the branch point. The result of the 2’ hydroxyl attack is the cleavage of the phosphodiester bond at the 5′ splice site, accompanied by the formation of a new 5′-2′ phosphodiester bond linking the first nucleotide of the intron with the special internal adenosine. As a result, the intron has been looped back on itself to create a lariat structure while the 5’ exon is released. In the second step of the splicing reaction, cleavage of the 3` splice site and ligation of the liberated exons occurs. A transesterification reaction is promoted by the 3`OH group from the end of the liberated upstream exon. This group attacks the phosphodiester bond at the 3` splice site, cleaving it and releasing the intron as a lariat structure. At the same time, the 3` end of the upstream exon joins the newly-formed 5` end of the downstream exon, completing the splicing process. The final products of the splicing reaction are the two joined exons and the intron, which is still in its lariat form (Figure 2) [Berget *et al.* 1977, Padgett *et al.* 1986, Sharp 2005].

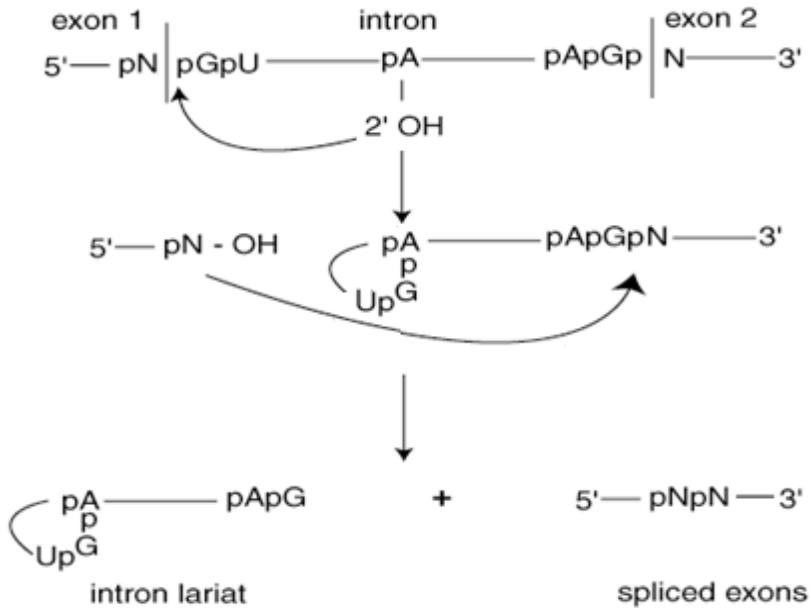


Figure 2 Two-step chemical mechanism for pre-mRNA splicing. Arrows represent first and second transesterification reaction, respectively.

The machinery which removes introns form a transcribed pre-mRNA is called a spliceosome. Two types of spliceosomes have been identified in eukaryotes, referred to as the major or U2 type and the minor or U12 type. In order to determine which type of spliceosome is required for intron splicing, we can discriminate two groups of introns: U2-type and U12-type introns. Generally, intron-flanking sequences were used to discriminate between these two groups. U2-type introns contain conserved GU and AG dinucleotides at their 5` and 3` ends, respectively [Reddy 2007; Reddy *et al.* 2013]. The second group (U12-type introns) was characterized by AU and AC at their 5` and 3` ends, respectively. However, the majority of U12-type introns contain GT-AG dinucleotides at their 5' and 3' ends, and a small number contain other noncanonical terminal dinucleotides [Lewandowska *et al.* 2004]. Compared to U2-type introns, U12-type introns lack a polypyrimidine tract and have a short distance between the branchpoint and the 3' splice site [Tarn & Steitz 1997]. For assembly and functioning of the spliceosome (which is a large ribonucleprotein [RNP] complex), approximately 100 proteins and five small nuclear RNAs (snRNAs) are required. For the major spliceosome, five types of snRNAs were identified and labelled U1, U2, U4, U5, and U6 sRNAs. These snRNAs are encoded as multicopy genes in the *A. thaliana* genome (14 genes for U1snRNA, 18 genes for

U2snRNA, 11 genes for U4snRNA, 14 genes for U5snRNA, and 13 genes for U6snRNA) [Wang & Brendel 2004]. Four types of snRNA genes are transcribed by RNA polymerase II, while U6 snRNA genes are transcribed by RNA polymerase III [Waibel & Filipowicz 1990]. U5 snRNA is used in both the major and minor spliceosomes, while U11, U12, U4atac, and U6atac replace the U1, U2, U4, and U6 snRNAs, respectively [Tarn & Steitz 1997]. SnRNAs are bound by many common proteins (Sm core proteins) and specific proteins identified in each snRNP. In *A. thaliana*, there are 15 Sm core proteins; seven of them were identified in U1, U2, U4, and U5 snRNPs, and seven (called Lsm – Like sm) are the counterparts of Sm proteins in the U6 snRNP. An additional Lsm1 protein cannot bind to snRNA [Will & Lührmann 2001]. The Sm domain of these proteins mediates protein interactions with other core proteins and with snRNP-specific proteins.

Splicing of a pre-mRNA that contains two exons and emphasizing the involvement of the snRNP in splicing is shown in Figure 3. From the very beginning of the splicing reaction, the assembly of a spliceosome is required. This assembly is started at the 5' splice site of the pre-mRNA [Reed 1996]. Recognition of the 5' splice site is based on RNA sequence complementarity between U1 snRNA (from the U1snRNP) and a few nucleotides at the boundary of the exon and intron [Kandels-Lewis & Seraphin 1993]. Next, U2 snRNP in the ATP-dependent process specifically binds to the branch-point region of the pre-mRNA intron [Barabino *et al.* 1990]. Subsequent to the binding of U2 snRNP, a tri-snRNP complex containing U4/U6 and U5 snRNP also associates in an ATP-dependent manner to the pre-mRNA. The U1 snRNP dissociates at this stage, and the 5' splice site intron sequence is bound by U6 snRNP after U4 snRNP destabilization. U4 and U6 snRNA (from the U4 and U6 snRNPs, respectively) enter the spliceosome paired with each other [Teigelkamp *et al.* 1994, Roy *et al.* 1995]. Transition to the second transesterification reaction requires action from some ATPases. The catalytic site of this reaction is created by U6 snRNA or U2 snRNA. The spliced RNA is released from the spliceosome through the action of a protein containing a helicase domain, and another putative RNA helicase is important for releasing the lariat intron [Company *et al.* 1991, Schwer & Gross 1998].

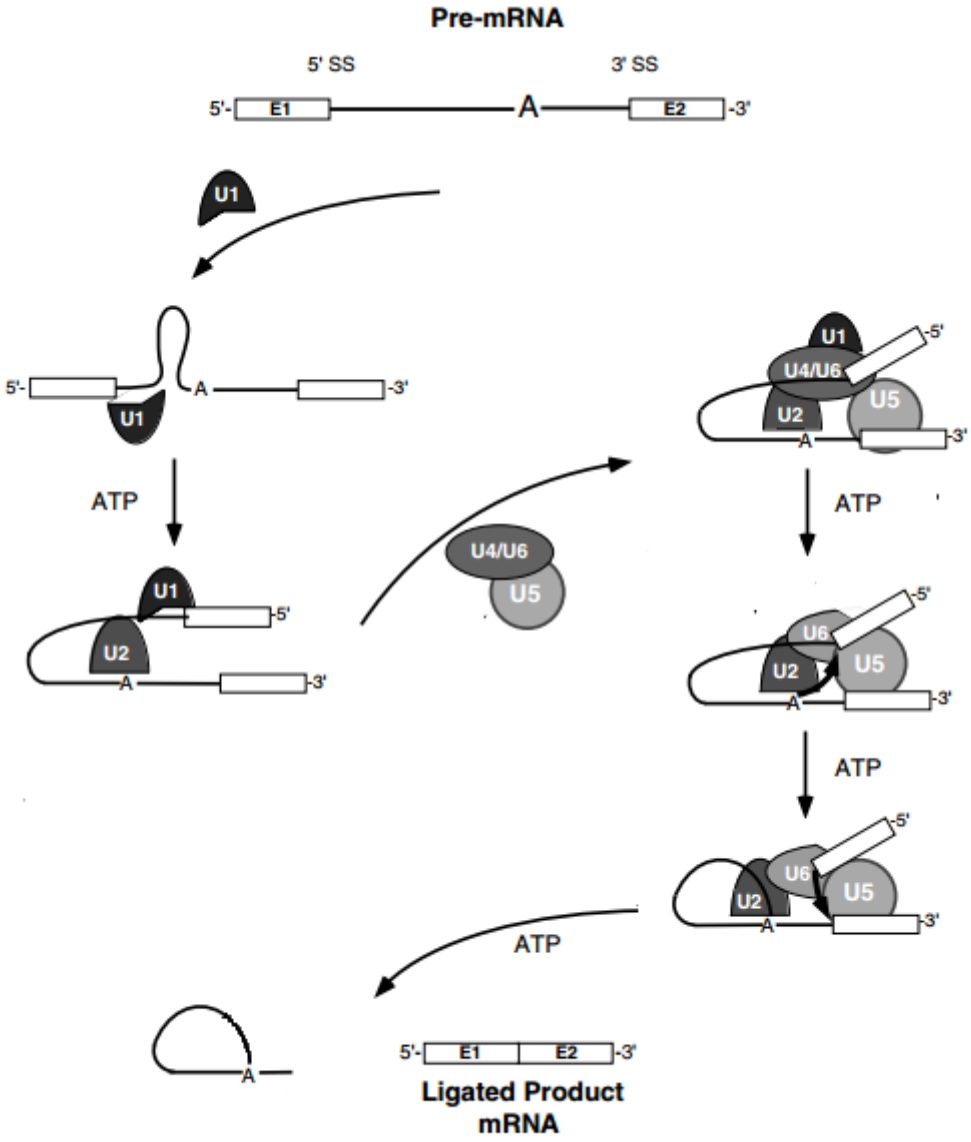


Figure 3 Overview of the roles of snRNPs in the processing of a pre-mRNA containing two exons and one intron into the ligated exon product and lariat intron.

4 Splicing of primary transcripts of microRNA genes in plants

Up till now, our knowledge about the regulatory role of splicing primary transcripts of microRNA genes in plants is limited. In the case of *MIR162a* and its host gene, the relationship between splicing and processing pri-miRNAs was described by John Brown in 2008. Nucleotide sequence coding for miR162a is located within the second intron of the At5g08185 gene (see Table 1 above). The primary transcript of this gene undergoes alternative splicing, resulting in the formation of six different mRNA isoforms. Three of them (including an unspliced transcript) may give rise to a mature miR162a.

Unfortunately, the detailed function of the three non-containing microRNA splice isoforms remains unknown. Schematic representation of alternative splicing of the At5g08185 gene is shown in Figure 4. Interestingly, miR162a targets *DCL1* mRNA for degradation. This example demonstrates the potential of competition between splicing and miRNA production in the fine tuning of regulation of gene expression [Hirsch *et al.* 2006; Brown *et al.* 2008].

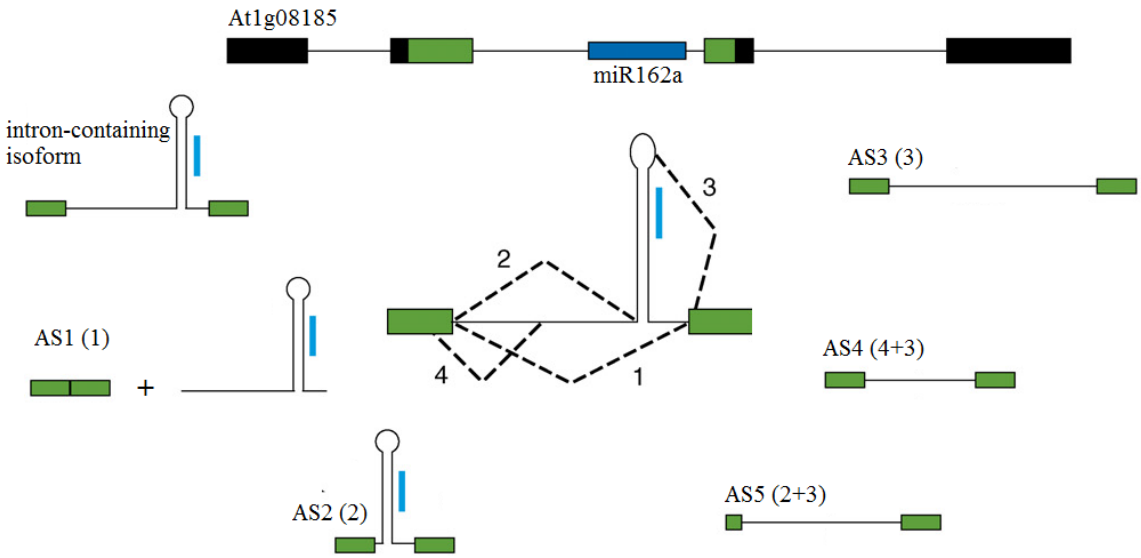


Figure 4 Schematic representation of alternative splicing of the At5g08185 gene, which is the host gene of miR162a. Only the intron-containing pre-mRNA, AS1, and AS2 contain the pri-miRNA and can give rise to a mature miR162a. Black boxes represent UTRs, green boxes represent exons, and solid lines represent introns [Brown *et al.* 2008, modified].

Another example of intronic microRNA is *A. thaliana MIR400*. Kang Yan and his colleagues proposed a regulatory model in which splicing can influence the fine-tuning expression of miR400 [Yan *et al.* 2012]. MiR400 is co-transcribed with its host gene, At1g32583, and is embedded in the first intron in the 5` UTR. In non-stressed conditions, the first intron in the 5` UTR is released from pre-mRNA by a spliceosome. According to the authors, this intron is further processed by a DCL1 to give rise to a mature miR400. On the other hand, the spliced isoform of the mRNA serves as a template for protein synthesis. Stress conditions induced by high temperatures lead to a decreased level of mature miR400s in *Arabidopsis* but do not affect host gene expression. Yan and his colleagues

showed that this decreased accumulation of miR400 is a consequence of alternative splicing of the miRNA-containing intron. Under heat treatment, a cryptic 5' splice site that is downstream of the miR400 sequence is selected. This selection leads to the excision of a 100 base pair fragment from the 3' end of the first intron and leaves miR400 in the transcript. According to the authors, this leads to decreased efficiency of processing of the transcript into mature microRNA. Schematic representation of a regulatory model of intronic miR400 biogenesis is presented in Figure 5. Additionally, in the transgenic plants overexpressing *MIR400* under heat-stress conditions, plants have a lower germination rate and decreased elongation of the hypocotyl when compared to wild-type plants. The obtained results demonstrate that a decreased level of miR400 under heat stress made the plants less sensitive to heat [Yan *et al.* 2012]. These observations clearly demonstrate that the accurate selection of splicing sites can be an important step in affecting microRNA expression.

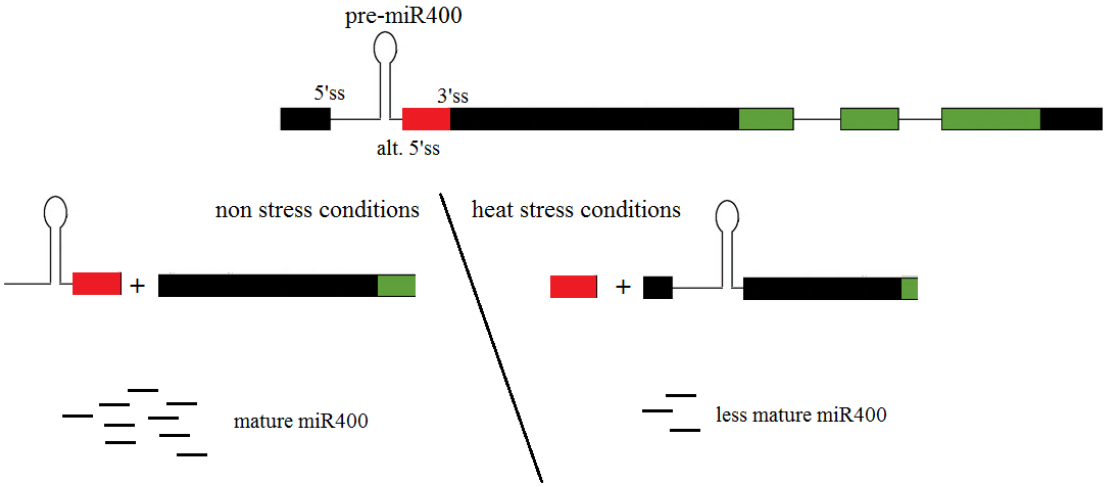


Figure 5 Schematic representation of a regulatory model of intronic miR400 biogenesis. Black boxes represent UTRs, green boxes represent exons, and red boxes represent alternatively spliced isoforms.

The two examples described above represent intronic microRNAs. However, the possible function or connection between the splicing and processing of *MIR* genes that are located in intergenic regions have not been investigated in details. It was thought that the region outside of the stem-loop is dispensable and lacks any regulatory role. However, we decided to investigate a connection between the dicing and splicing of plant pri-miRNAs in detail.

The aim of the work

The presence of introns in plant *MIR* genes is a common phenomenon. It is known that splicing of so called intronic *MIR* genes can modulate expression of their mature microRNA forms. However, the connection between splicing and dicing of *MIR* genes which are independent transcriptional units remains unknown. One of the aims of this dissertation was to get deeper insight into this connection.

An involvement of *A. thaliana* HYL1 protein in the maturation of miRNA primary precursors is unquestionable. It was shown that pri-miRNAs accumulate in *hyl1-2* mutants as compared to wild-type plants. Moreover, in the *hyl1-2* mutant also an unspliced isoform of several pri-miRNAs accumulates. Additionally, it is well known that intron removal by spliceosme is a co-transcriptional event. This observation suggests another HYL1 role. The second aim of this dissertation was to investigate an involvement of HYL1 protein in the transcription of *MIR* genes.

Experimental works included:

1. Comparison of the maturation of intron-containing and intron-less maturation pri-miRNAs.
2. Analyses of the splice sites requirements witin pri-miRNAs.
3. Studies on the splicing of the pri-miRNA and fuction of its mature microRNA form.
4. Analyses of the HYL1 influene on GUS reporters genes.
5. Characteristics of RNA POL II distribution in wild-type and *hyl1-2* plants.

MATERIALS

1. Bacterial strains

Escherichia coli DH5α competent cells were used for cloning and plasmid amplification,

Agrobacterium tumefaciens AGL1 cells were used for transformation of *Arabidopsis thaliana* plants using the *floral dip* method and for transient expression experiments in *Nicotiana benthamiania,*

Pseudomonas syringae DC3000 cells were used for pathogen-infection experiments.

2. Plant material

The experiments were performed using *Arabidopsis thaliana* (L.) Heynh., Columbia-0 and *Nicotiana benthamiana* wild-type plants. Additionally, *Arabidopsis thaliana* homozygous T-DNA insertion lines were used: *mir163-2* (SALK_0034556), *sr34-1* (SALK_106067), *rs31-1* (SALK_021332), *rs2z33-1* (GABI_180D12), and *scl30a-1* (SALK_095431) [Bielewicz *et al.* 2013, *hyl1-2* (SALK_064863) and *hen1-5* (SALK_049197) [Vazquez *et al.* 2004a], *cbp20* [Papp *et al.* 2004], *cbp80 (abh1)* [Hugouvieux *et al.* 2001], and *cbp20×cbp80 (abh1)* double mutant (obtained by Paulina Piontek from the Department of Gene Expression, Adam Mickiewicz University). Point mutation mutant *cpl1-7* was obtained from Detlef Weigel's lab [Manavella *et al.* 2012]. GUS reporter lines (TIR1p:GUS, AFB1p:GUS, AFB2p:GUS, AFB3p:GUS, p393A:GUS, p393B:GUS) were obtained from Marke Estelle's lab from the Division of Biological Sciences, the University of California, San Diego.

3. Oligonucleotides

Designed primers were purchased in a lyophilized form from (i) Genomed (Warsaw, Poland) or (ii) Sigma Aldrich (St. Louis, Missouri, USA).

Primers used for construct preparation:

Name	Nucleotide sequence 5'→ 3'	cDNA/gene fragment amplified using the primer pair	Constructs prepared using the amplified fragment
A01	TTGCGGCCGCTCGTGAATCTTTGTTTCCTC	native promoter + *MIR163*	IVSwt
A02	TTGGCGCGCCCAAGCGTCCAGACTTCAG		

A03	TTGCGGCCGCTCGTGAATCTTTGTTTCCT	native promoter + 1st exon of *MIR163*	ΔIVS
A04	GAGGAAACAAAAAATTTCCGTTATCTCTTTTCATC		
A05	GATGAAAAGAGATAACGGAAATTTTTTGTTTCCTC	2nd exon of *MIR163*	
A06	TTGGCGCGCCCAAGCGTCCAGACTTCAG		
A07	GATGAAAAGAGATTTTTTTAGTCATGCACATG	mutagenesis of IVSwt*MIR163*	IVSΔ5`ss; IVSmut
A08	CATGTGCATGACTAAAAAAATCTCTTTTCATC		
A09	GTCTAATGATTTTTTTTTAATTTTTTGTTTC	mutagenesis of IVSwt*MIR163*	IVSΔ3`ss; IVSmut
A10	GAAACAAAAAATTAAAAAAAAATCATTAGAC		

Name	Nucleotide sequence 5'→ 3'	cDNA/gene fragment amplified using the primer pair	Constructs prepared using the amplified fragment
A11	TTGCGGCCGCCACTTATCTCTAACTCATCC	*MIR161*	IVSwt
A12	TTGGCGCGCCTGTCTTCTTCTTCTCTTGTG		
A13	TTGCGGCCGCCACTTATCTCTAACTCATCC	1st exon of *MIR161*	ΔIVS
A14	CTTTTAAAAACTTTCTCGCATCACAATTTCAATGCTTTTCC		
A15	GGAAAAGCATTGAAATTGTGATGCGAGAAAGTTTTTAAAA	2nd exon of *MIR161*	
A16	TTGGCGCGCCTGTCTTCTTCTTCTCTTGTG		
A17	CATTGAAATTTTTTTTGGAGATGGATATG	mutagenesis of IVSwt *MIR161*	IVSΔ5`ss
A18	CATATCCATCTCCAAAAAAAATTTCAATG		
A19	GTTCAATTGTTATTTTTTTTTTTCGAGAAAGTTTTTAAAAG	mutagenesis of IVSwt *MIR161*	IVSΔ3`ss
A20	CTTTTAAAAACTTTCTCGAAAAAAAAAAATAACAATTGAAC		
A21	TTAAGCTTGCTTCATCACCTTCCCACAT	promoter region of ACT2	pMDC-pACT2
A22	TTAAGCTTTTTATGGGTGGATTGTGGTG		
A23	TTAAGCTTAGCTTGAATCTCCCTCGTGA	promoter region of GAPDH	pMDC-pGAPDH
A24	TTGGCGCGCCTTTGCGAAATTGAGATCGAGAGAG		

Primers used for RT-PCR, qPCR analyses:

Name	Nucleotide sequence 5'→ 3'	cDNA/gene fragment amplified using the primer pair	Experiments in which the primer pair was used

B01	GAGAGTGAGAAAAATAAAGAG	pri-miR163	RT-PCR
B02	AGGATGTTGACACGTGTAAAC		
B03	AGTACCTTAGATAAACCGACCAAAACC	pri-miR163 (1st exon)	qPCR & poly(A) sites usage ratio
B04	AACCGGGAACTCCAGCACTT		
B05	GAATGCAAATGGTTGTGGAA	pri-miR163 (2nd exon)	
B06	GGGCCACTAAAGCCCTTAAA		
B07	GGCATCAGATTTCACCTTTTTC	pri-miR161	qPCR
B08	CAAATGATGCAATCTCAAACAAA		
B09	ATTTCGGCTCCAACAATGTC	HPTII	qPCR
B10	GATGTTGGCGACCTCGTATT		
B11	TGCCGAAGCTTTGATCAGTA	pri-miR161	RT-PCR
B12	TCAAATGATGCAATCTCAAACA		
B13	ACATTGTTGGAGCCGAAATC	HPTII	RT-PCR
B14	GTGCTTGACATTGGGGAGTT		
B15	GGGATCTCGGACATCTTAACTCCGTCTCTC	At1g66690	5` RACE
B16	CCCCGGAGGGAGTGTCTGGA	At1g66690	qPCR
B17	GGCGATCTCCGCCAACCTCG		

Oligonucleotides used as probes in Northern blot hybridizations:

Name	Nucleotide sequence 5'→ 3'	Detection of miRNA
C01	ATCGAAGTTCCAAGTCCTCTTCAA	miR163
C02	CGTGATATTGGCACGGCTCAA	miR171
C03	TAGTCACTTTCAATGCATTGA	miR161
C04	GGATCAATGCGATCCCTTTGGA	miR393
C05	GGGGAATGAAGCCTGGTCCGA	miR166
C06	TAGATCATGCTGGCAACTTCA	miR167m
C07	CGTGATATTGGCACGGCTCAA	miR171
C08	ATGCAGCATCATCAAGATTCT	miR172

Primers used for ChIP experiments:

Name	Nucleotide sequence 5'→ 3'	gene fragment amplified using the primer pair
D01	CATTTTACCCCACCATTTGC	promoter region of *MIR*393A
	CGTGAAGAATGTGTGCGTTC	
D02	CAATGCAAATTTGTGCGTGT	promoter region of *MIR*393A

	GGATCCAACGTATCTCTGTCG	
D03	TTGCTGTGATAGAGCGTGTTTT	promoter region of *MIR*393A
	TGATTTTGCGGTTGACACAT	
D04	GCTTGGTTTTGGATCATGCTATCTCTT	body-gene region of *MIR*393A
	CGGAATCTTAAAGCCAGCAAAG	
D05	ATTTGCAAACATGGATACGC	3` downstream region of *MIR*393A
	GGTGTTTTGTGCTTTTCCTTC	
D06	TGGTTTACGAAGCTGCATGT	promoter region of *MIR*393B
	GAGAGAAGGAGAATATCGATGACA	
D07	CCGTAGAAGCAATAGGCAGAA	promoter region of *MIR*393B
	CGGGAAACTTGCTTTTTGAT	
D08	TGCATAGCCTTGCAACAAAA	promoter region of *MIR*393B
	CCAACCACCATCAATTCATTT	
D09	TCTAGCACGCACAGAGAGGA	body-gene region of *MIR*393B
	TTCCTCCACCAACTCAACATC	
D10	CAAGCTACTACGCGCAACTG	3` downstream region of *MIR*393B
	ACAGCCCATCTTAGGGTTTT	
D19	TGGCTCTCTCCACTACTCAA	promoter region of *MIR166A*
	GACAACAGTCCCCTCAAAA	
D20	CGACCCTTAAACTCTCCATAA	promoter region of *MIR167A*
	ACTTCACCGTAGCAGATCAA	
D21	TGCTTTGGTAGTAGATGAGGTT	promoter region of *MIR171Aa*
	CGTGTGTGGTCAGGTAAGAT	
D22	TATTAAGGACTTGTAGGACTCA	promoter region of *MIR172B*
	TAATAGTACGTACACATAAATGG	
D23	AGTTCAATGGAGAGATGTCGAAATATG	untranscribed region of pol II-CI
	AAGAGGAAAAGAAAGAGATGGAGAGA	

4. Vectors

For cloning and construct preparation, the following vectors were used:

- pGEM T-Easy plasmid vector (Promega)
- pENTR/D-TOPO (Life Technologies)
- pMDC32, pMDC99 and pMDC123 [Curtis & Grossniklaus, 2003]

5. Nucleic acids molecular weight markers

1- 0'GeneRuler™ 100 bpPlus DNA Ladder Thermo Scientific

2 - 0'GeneRuler™ 1kb Plus DNA Ladder Thermo Scientific

6. Enzymes

Name	Company
Taq DNA Polymerase 5U/µl	Thermo Scientific
Pfu DNA Polymerase 2.5U/µl	Thermo Scientific
T4 DNA Ligase 5U/µl	Thermo Scientific
T4 Polynucleotide Kinase 10U/µl	Roche
SuperScript III Reverse Transcriptase 200U/µl	Life Technologies
FastDigest NotI	Thermo Scientific
FastDigest SgsI (AscI)	Thermo Scientific
FastDigest PvuII	Thermo Scientific

7. Kits

Name	Purpose	Company
QIAquick Gel Extraction	Extraction and purification of DNA fragments from gel	Qiagen
GenElute Gel Extraction	Extraction and purification of DNA fragments from gel	Sigma Aldrich
QIAquick PCR Purification	Purification of DNA after enzymatic reactions	Qiagen
GenElute PCR Clean-Up	Purification of DNA after enzymatic reactions	Sigma Aldrich
GenElute Plasmid Miniprep	Isolation and purification of plasmids from bacteria	Sigma Aldrich
TURBO DNA-free	DNase treatment of RNA	Life Technologies
Gateway LR Clonase II Enzyme Mix	Cloning to destination vectors	Life Technologies
2xPower SYBR Green PCR Master Mix	Real-time PCR	Life Technologies
IllustraMicroSpin G-25 Columns	Purification of radiolabeled oligonucleotides	GE Healthcare
SMART™ RACE cDNA Amplification	Rapid amplification of cDNA ends	Clontech

QuikChange II XL	Site-Directed Mutagenesis	Stratagene

8. Reagents

ACE (Liquid bleach)	Procter & Gamble
Acrylamide	Sigma
Agar, plant cell culture tested	Sigma
Agar, microbiological	Becton, Dickinson
Agarose	Prona
APS (Ammonium peroxydisulfate)	Sigma
Bactotryptone	Becton, Dickinson
Bacto yeast extract	Becton, Dickinson
Boric acid	Sigma
Bromophenol blue	Sigma
Chloroform	Chempur
DEPC (Diethyl pyrocarbonate)	Sigma
EDTA (Ethylenediaminetetraacetic acid)	Sigma
Ethanol	Chempur
Ethidium bromide	Sigma
Formamide	Chempur
Ferrocyanide	Sigma
Ferricyanide	Sigma
Glufosinate ammonium (BASTA)	Bayer
Glycerol	Sigma
Hydrochloric acid	Chempur
Hydroxy peroxide, 30%	Chempur
Isopropanol	Polskie Odzynniki Chemiczne
MS (Murashige & Skoog) medium, powdered	Duchefa Biochemie
N, N'-methylene-bis-acrylamide	Sigma
Phenol, Tris-saturated	Roth
SDS (Sodium dodecyl sulfate)	MP Biomedicals
Sodium citrate, tribasic dihydrate	Polskie Odczynniki Chemiczne
Sucrose	Sigma

TEMED (N,N,N',N'-Tetramethylethylenediamine)	Merck
Tris base (Tris(hydroxymethyl)aminomethane)	Sigma
TRIzol reagent	Life Technologies
Urea	Sigma
X-Gal (5-bromo-4-chloro-3-indolyl-D-galactopyranoside)	Sigma

9. Buffers and solutions

9.1. Bacterial culture

LB (Luria-Bertani) liquid medium

1% bactotryptone

0.5% bacto yeast extract

1% NaCl

autoclaved (121˚C, 1 atm., 20 minutes)

LB (Luria-Bertani) solid medium

1% bactotryptone

0.5% bacto yeast extract

1% NaCl

1.5% microbiological agar

autoclaved (121˚C, 1 atm., 20 minutes)

After autoclaving and cooling to approx. 50˚C, antibiotics were added to the medium. The medium was poured into sterile Petri dishes (Ø 9 cm), parafilmed, and kept upside-down at 4˚C until use.

YEB liquid medium

0.5% beef extract

0.1% yeast extract

0.5% peptone

0.5% sucrose

autoclaved (121˚C, 1 atm., 20 minutes)

After autoclaving, a sterile $MgCl_2$ water solution was added to the final concentration of 10 mM of the medium.

antiobiotic solutions

ampicillin 50 mg/ml, dissolved in water

kanamicin 50 mg/ml, dissolved in water

rifampicin 50 mg/ml, dissolved in DMSO

X-Gal solution

X-Gal 20 mg/ml, dissolved in N,N-dimethyloformamide

9.2. Plant growth

Seed sterilization solution

70% ethanol

1.2% sodium hypochlorite

Plant culture soil

Jiffy-7 42 mm soil (Jiffy International AS, Norway)

'Podłoże warzywne' (Holas, Poland)

Plant culture plates (half -strength MS medium)

0.22% MS (Murashige & Skoog) medium (DuchefaBiochemie)

1.5% sucrose

0.8% agar

autoclaved (121˚C, 1 atm., 20 minutes)

The medium was poured into sterile Petri dishes (Ø 9 cm), parafilmed, and kept upside-down at 4˚C until use.

9.3 Nucleic acid analyses

RNase-free water

0.05% DEPC

prepared using milliQ water (Millipore)

incubated overnight at 37˚C

autoclaved (121˚C, 1 atm., 40 minutes)

10x TBE buffer

890 mMTris base

890 mM boric acid
20 mM EDTA
autoclaved (121˚C, 1 atm., 20 minutes)

Agarose gel
1-2% agarose
1x TBE buffer
After boiling and cooling to approx. 50˚C, an ethidium-bromide solution was added to the final concentration of 0.05 mg/100 ml.

RNA loading buffer
10 mM Tris·HCl pH 7.5
2.5 mM EDTA
95% formamide
0.01% xylencyanol
0.01% bromophenol blue
The RNA loading buffer was prepared using RNase-free (DEPC-treated) water.

DNA loading buffer (2x HSE)
4 M urea
50% sucrose
50 mM EDTA
0.1% xylencyanol
0.1% bromophenol blue

10% SDS solution
10% Sodium dodecyl sulfate
filtrated (Ø 0.22 µm, Millipore)

10% APS solution
10% Ammonium peroxydisulfate
filtrated (Ø 0.22 µm, Millipore)
The APS stock solution was distributed to 1.5-ml tubes and stored at -20˚C.

SSC (20x)

3 M NaCl

0.34 M sodium citrate

autoclaved (121˚C, 1 atm., 20 minutes)

Hybridization buffer

375 mM Na_2HPO_4

125 mM NaH_2PO_4

1% SDS

The hybridization buffer was prepared using RNase-free (DEPC-treated) water.

9.4 GUS staining

1 M NaPi (pH 7.0)

57.7 ml 1 M $Na_2 HPO_4$

42.3 ml 1 M $NaH_2 PO_4$

X-Gluc solution (for 100 ml)

100 mM NaPi

5 mM K-Ferr**o**cyanide

5 mM K-Ferr**i**cyanide

100 mg X-Gluc

9.5 Chromatin Immunoprecipitation

Nuclei isolation buffer I

10 mM Tris HCl pH8

10 mM $MgCl_2$

400 mM sucrose

0.035% β-mercaptoethanol*

1 mM PMSF*

* Added just before use. Stored in 4˚C.

Nuclei isolation buffer II

10 mM Tris HCl pH8

10 mM $MgCl_2$

400 mM sucrose

1% Triton X-100

0.035% β-mercaptoethanol*

1 mM PMSF*

1 tab/ml cOmplete EDTA-free*

0.004 mg/ml Pepstatin A*

* Added just before use. Stored in 4˚C.

Nuclei isolation buffer III

10 mM Tris HCl pH8

2 mM $MgCl_2$

1.7 M sucrose

0.15% Triton X-100

0.035% β-mercaptoethanol*

1 mM PMSF*

1 tab/ml cOmplete EDTA-free*

0.004 mg/ml Pepstatin A*

* Added just before use. Stored in 4˚C.

HEPES

0.5 M HEPES

pH was adjusted to 7.5 using NaOH

Sonic buffer

10 mM sodium phosphate buffer

100 mM NaCl

0.5% sarkosyl

10 mM EDTA

1 mM PMSF*

1 tab/ml cOmplete EDTA-free*

0.004 mg/ml Pepstatin A*

* Added just before use. Stored in 4˚C.

Nuclei lysis buffer

50mM Tris HCl pH8

10mM EDTA

1% SDS

1 mM PMSF*

1 tab/ml cOmplete EDTA-free*

0.004 mg/ml Pepstatin A*

* Added just before use. Stored in 4˚C.

IP buffer

50 mM HEPES pH 7.5

150 mM NaCl

10 μM $ZNSO_4$

1% Triton X-100

0.05% SDS

1 mM PMSF*

1 tab/ml cOmplete EDTA-free*

0.004 mg/ml Pepstatin A*

* Added just before use. Stored in 4˚C.

ChIP dilution buffer

16.7 mM Tris HCl pH8

1.2 mM EDTA

1.1% Triton X-100

167 mM NaCl

1 mM PMSF*

1 tab/ml cOmplete EDTA-free*

0.004 mg/ml Pepstatin A*

* Added just before use. Stored in 4˚C.

Low salt buffer

20 mM Tris HCl pH8

2 mM EDTA

1% Triton X-100

150 mM NaCl

0.1% SDS

High salt buffer

20 mM Tris HCl pH8

2 mM EDTA

1% Triton X-100

500 mM NaCl

0.1% SDS

LiCl buffer

10 mM Tris HCl pH8

2 mM EDTA

250 mM LiCl

1% NP-40

1% sodium deoxycholate

TE buffer

10 mM Tris HCl pH8

1 mM EDTA

Elution buffer

10 mM Tris HCl pH8

1 mM EDTA

1% SDS

METHODS

1. Plant growth

1.1 Seed sterilization

Approximately 100 µl of *A. thaliana* seeds were put in a 1.5 ml eppendorf tube. Next, 1 ml of sterilization solution was added to the seeds, and the tubes were mixed by inverting them for 10 minutes. After incubation time, the seeds were washed three times with 1 ml of 100% ethanol and dried overnight in a flow hood.

1.2 Plant growth conditions

• Plant culture in soil

A. thaliana plants were grown in 'Jiffy-7 42 mm' soil (Jiffy International AS, Norway) in an MLR-350H Versatile Environmental Test Chamber (Sanyo, UK) with a 16-h day (approx. 150-200µE/m^2), a constant temperature of 22°C, and 70% humidity.

N. benthamiana plants were grown in 'Podłoże warzywne' soil (Holas, Poland) in an MLR-350H Versatile Environmental Test Chamber (Sanyo,UK) with a 16-h day (approx. 150-200 µE/m^2), a constant temperature of 25°C, and 60% humidity.

• Plant culture on plates

A. thaliana plants were grown on a half-strength MS solid medium. Sterilized seeds were sown on selection plates under sterile conditions (approx. 80 seeds per 10-sq-cm plate). Plates were parafilmed and placed horizontally in a growth chamber (22°C with a 16-h day).

1.3 Plant transformation

1.3.1 Transformation of *Arabidopsis thaliana*

Transformation of plants was done using the *floral dip* method. The protocol for transformation was based on a publication by Steve Clough and Andrew Bent from the University of Illinois at Urbana-Champaign [Clough & Bent, 1998].

1. *A. thaliana* plants were grown in Jiffy pots in long-day conditions. The culture was started approximately four weeks before transformation.

2. The *A. tumefaciens* strain carrying the gene of interest on a binary vector were grown in a large (250 ml) liquid culture in 28^{o}C in an LB medium with antibiotics (rif: 100 μg/ml, kan: 50 μg/ml).

3. A culture of *Agrobacterium* was centrifuged for 20 minutes at 3000g, and the cells were resuspended to OD_{600} = 0.8 in a 5% sucrose solution. Before dipping, Silwet L-77 was added to a concentration of 0.02%.

4. During the day of transformation, all siliques were removed from the plants.

5. The prepared plants were dipped in the *Agrobacterium* solution for 2 to 3 seconds, with gentle agitation.

6. Dipped plants were placed under a cover for 24 hours at room temperature and in total darkness.

7. Next, the plants were grown normally until their seeds became mature.

8. Dry seeds were harvested and placed in envelopes, ready for selection.

1.3.2 Transformation of *Nicotiana benthamiana*

1. *Nicotiana benthamiana* plants were grown for 6 weeks at 25°C in a Sanyo MLR-351H chamber (16/8 h light/dark; 50% humidity; 150–200 μmol m^{-2} s^{-1}).

2. *A. tumefaciens* strain carrying the gene of interest on a binary vector were grown in a large (250 ml) liquid culture in 28^{o}C in an LB medium with antibiotics (rif: 100 μg/ml, kan: 50 μg/ml).

3. A culture of *Agrobacterium* was centrifuged for 20 minutes at 3000g, and the cells were resuspended to OD_{600} = 0.6 in 10 mM MES pH 5.6, 10 mM $MgCl_2$

5. The prepared *Agrobacterium* solution was infiltrated into leaves using a syringe.

6. *Nicotiana* plants were grown for additional 72 hours after transformation.

7. Next, tissue from the infiltrated leaves was collected.

1.4 GUS staining

Seedlings were placed in a sterile 14 ml plastic tube and covered with an X-Gluc stock solution. Next, the tube was incubated at 37 °C for 24h. The X-Glux solution was removed, and the seedlings were placed in 70 % (v/v) EtOH. The ethanol solution was changed until the leaves were no longer green. At the end, the ethanol was removed, and the seedlings were placed in a 50% glycerol solution.

2. Bacterial growth

2.1 Bacterial transformation

•Transformation of *Escherichia coli*

Bacterial chemi-competent cells (100 µl) were thawed on ice (approximately 15 minutes). The ligation mixture was added to bacterial competent cells and mixed gently with the end of a tip. The sample was incubated at 4˚C for 15 minutes. Bacterial cells were incubated for exactly 2 minutes at 42˚C and immediately chilled on ice for additional 2 minutes. 900 µl of an LB liquid medium (pre-warmed to 37˚C) was added to the transformed bacterial cell suspension. After mixing by inverting the tube, the sample was incubated for 1 hour at 37˚C with gentle shaking (Thermomixer Comfort, Eppendorf).

•Transformation of *A. tumefciens*

Bacterial electro-competent cells were thawed on ice (approximately 15 minutes). The ligation mixture was desalted for 30 minutes using a nitrocellulose 0.025 µm filter (Millipore, USA). Next, the solution was added to bacterial competent cells and mixed gently with the end of a tip. Competent cells were placed in a pre-chilled 0.1-cm-gap electroporation cuvette (Bio-Rad, USA). A pulse of 2.5 kV (129 ohm resistance) was given, and 1 ml of a liquid LB medium was immediately added. Bacteria were gently mixed and transferred to a 2 ml eppendorf tube and incubated at 28°C for 1 hour with gentle shaking.

2.2 Selection of bacterial transformants

200 µl of transformed bacterial cell suspension (along with 30 µl of X-Gal solution after transformation with pGEM-T Easy plasmid) were spread on agar selection plates with proper antibiotics for each selection. Plates with bacteria were incubated upside-down at 37˚C for 10-15 hours. Bacteria-that formed white colonies were grown in a liquid LB medium with proper antibiotics.

2.3 Bacterial liquid culture

To amplify the transformed bacterial cells, a single colony was picked with a sterile tip and transferred to a sterile glass flask with 5 ml of an LB liquid medium with proper antibiotics (pre-warmed to 37˚C). Bacteria were grown at 37˚C with shaking for 10-15 hours. The obtained bacterial liquid culture was used for the isolation of plasmid DNA.

2.4 *Pseudomonas syringae* bacterial treatment

P. syringae DC3000 cells were grown in a liquid LB medium for 10-15 hour at 28°C. The culture was centrifuged for 20 minutes at 3000g, and the cells were resuspended to OD_{600} = 0.2 in a 10 mM $MgCl_2$ solution. Two-week-old plants grown in soil were sprayed with a prepared bacteria solution. 24 and 72 hours after treatment, aerial parts of the sprayed plants were collected.

3. Methods used during work with nucleic acids

3.1 RNA isolation

Total RNA was isolated from the grinded in liquid nitrogen plant material. Isolations were carried out using a TRIzol reagent. Briefly, 1 ml of the TRIzol reagent was added to 100 - 150 mg of powdered material and mixed by vigorous vortexing to eliminate visible clumps. Samples were incubated in a horizontal position for 5 minutes at room temperature. After incubation, 200 µl of chloroform was added, and the samples were mixed by vortexing. After incubation in a horizontal position for 3 minutes, the samples were centrifuged for 10 minutes at 4ºC at 12,000g. The obtained supernatant was transferred into new tubes. The samples were mixed with an equal volume of isopropanol, mixed briefly by vortexing, and incubated for 20 minutes at room temperature. After incubation, the samples were centrifuged for 20 minutes at 4ºC at 12,000g, and the supernatant was discarded. A pellet of RNA visible at the bottom of the tube was washed by adding 1 ml of chilled 70% ethanol. The samples were centrifuged for 5 minutes at 4ºC at 12,000g. After discarding the supernatant, the RNA pellet was air-dried at room temperature and dissolved in RNase-free water. Sample concentration was checked using a NanoDrop spectrophotometer.

3.2 DNase digestion

RNA was digested using DNase I (TURBO DNA-free kit, Life Technologies) according to the manufacturer's protocol.

3.3 First strand cDNA synthesis

Reverse transcriptase reactions were performed using Oligo(dT)18 Primer (Thermo Scientific) and SuperScript III Reverse Transcriptase (Life Technologies) according to the manufacturer's protocol.

3.4 Northern hybridization

3.4.1 RNA electrophoresis in PAA

RNA electrophoresis was performed in 15% (19:1) polyacrylamide gel containing 8M urea. Glasses were washed with 5% SDS and then cleaned with ethanol. Gel was poured between the glasses immediately after the addition of TEMED and APS. Before loading the RNA samples, the gel was run for 30 minutes at 300V (1XTBE buffer was pre-heated to 65°C). The RNA samples (approx. 15 µg) were mixed with an equal volume of formamide-loading dye, heated for 2 min at 95°C, and loaded onto the gel. Electrophoresis was performed at 300V until the bromophenol blue dye migrated to the bottom of the gel.

3.4.2 Transfer of RNA from gel to membrane

Four pieces of blotting paper and a piece of Amersham Hybond NX nylon membrane was cut slightly larger than the gel surface. All of the prepared pieces were soaked in 0.5X TBE. Next, two pieces of blotting paper were stacked on the anode platform of the transfer cell (Pierce® Fast Semi-Dry Blotter, Thermo Scientific). Air bubbles were removed by rolling a glass pipet across the surface. The membrane was placed on top of the blotting paper. Gel was carefully transferred from the glass plate to the top of the membrane, and the air bubbles were squeezed out. Two more pieces of blotting paper were then placed on the gel, and again, the air bubbles were squeezed out. The cathode platform was assembled, and the transfer was carried out for 1 hour at 10V.

After the transfer, the transfer cell was disassembled. The blotting paper and gel were removed. The membrane was placed RNA-side-up in a UV-light box (Stratagene® UV Stratalinker). The RNA was covalently attached to the membrane by applying UV light energy at a level of 120 kJ/cm^2.

3.4.3 Radiolabeling of oligonucleotides

Designed DNA oligonucleotides were perfectly complementary to a given small RNA to be detected. The 5`-end labeling reaction was set up by combining the following reagents:

2 µl of 10 µM probe (DNA oligonucleotide)

5 µl of [γ^{32}P] ATP (6000 Ci/mmol)

2 µl of T4 polynucleotide kinase buffer

1 µl of T4 polynucleotide kinase 10U/µl (Roche)

$H2O_{DEPC}$ to a final volume of 20 µl

The reaction mixture was incubated for 45 min at 37°C. Next, unincorporated [γ^{32}P]ATPs were removed from the samples using Illustra MicroSpin G-25 columns, according to the manufacturer's protocol.

3.4.4 Hybridization of radio-labeled oligonucleotides with RNA

The membrane was placed RNA-side-up in a hybridization bottle, and 5 ml of a pre-warmed (42°C) hybridization buffer was added. The bottle was placed in a hybridization oven and incubated (with rotation) for 30 min at 42°C. After incubation, the hybridization solution was replaced with a fresh hybridization solution. The reaction mixture with the radiolabeled oligonucleotide was added to the bottle. Next, the bottle was again placed in the hybridization oven and incubated (with rotation) for 10-15 hours at 42°C. After overnight incubation, the hybridization solution was discarded, and the membrane was briefly washed two times with 2X SSC for 15 minutes.

3.4.5 Detection of signal after hybridization

Hybridization signals were visualized by phosphor imaging using Image an FLA5100 from Fuji. Hybridization results were analyzed using Multi Gauge v2.2 software.

3.5 PCR reactions

3.5.1 PCR reactions using DreamTaq DNA polymerase

For standard PCR reactions, the DreamTaq DNA polymerase from the Fermentas company was used. The reaction mixture contained 1X DreamTaq DNA Polymerase Buffer with MgCl2, 0.2 mM each dNTP, 0.5 μM forward primer, 0.5 μM reverse primer, and DreamTaq DNA Polymerase 0.05 U/μl. Genomic DNA or cDNA was used as a template. The final volume of the reaction was 20 μl. Negative controls contained water as a template.

The thermal profile was set as presented below:

PCR step	Temperature	Time	Number of repeats
template denaturation	95°C	3 minutes	1
template denaturation primer	95°C 55-65°C	30 seconds 30 seconds	25-35

annealing elongation	72°C	1-5 minutes	
final elongation	72°C	1-15 minutes	1
cooling down	8°C	15 minutes	1

3.5.2 PCR reactions using *Pfu* DNA polymerase

For construct preparation, the *Pfu* DNA polymerase from the Fementas company was used. The reaction mixture contained 1x *Pfu* DNA Polymerase Buffer with Mg_2SO_4, 0.2 mM each dNTP, 0.5 μM forward primer, 0.5 μM reverse primer, and *Pfu* DNA Polymerase 0.05 U/μl. Genomic DNA or cDNA was used as a template. The final volume of the reaction was 50 μl. Negative controls contained water as a template.

The thermal profile was set the same as the DreamTaq DNA polymerase.

3.5.3 Quantitative real-time PCR

Real-time PCR reactions were performed using the 7900HT FastReal-Time PCR System (Life Technologies) and SYBR Green to monitor dsDNA synthesis. The reaction mixture contained 5 μl of 2X Power SYBR Green PCR MasterMix (Life Technologies), cDNA, and gene-specific primers (200 nM each) in a final volume of 10 μl. After each real-time PCR run, dissociation-curve analyses were performed.

The thermal profile was set as presented below:

Real-time PCR step	Temperature	Time	Number of repeats
template denaturation & polymerase activation	95°C	10 minutes	1
template denaturation primer annealing & elongation	95°C 60°C	15 seconds 1 minute	40
denatruation	95°C	15 seconds	1
denatruation	60°C rising to 95°C	30 minutes	1

The obtained results were analyzed using SDS 2.3 and Excel software.

3.6. Genomic DNA isolation

One leaf from an approximately two-week-old plant was placed in a 1.5 ml eppendorf tube. The leaf was ground in liquid nitrogen with a plastic stick for 15 seconds, and 400 µl of DNA extraction buffer was immediately added. Next, the sample was mixed by vortexing for 30 seconds, followed by incubation for 1 hour at room temperature. The tube was centrifuged for 15 minutes at 4°C at 12,000g. 300 µl of supernatant was transferred into a new tube. The sample was mixed with 300 µl of isopropanol, mixed briefly by vortexing, and incubated for 15 minutes at room temperature. After incubation, the sample was centrifuged for 15 minutes at 4ºC at 12,000g, and the supernatant was discarded. A pellet of DNA visible at the bottom of the tube was washed by adding 200 µl of 70% ethanol. The sample was centrifuged again for 5 minutes at 4ºC at 12,000g. After discarding the supernatant, the DNA pellet was dissolved in 100 µl of water.

3.7 Agarose gel electrophoresis

For running agarose-gel electrophoresis, a small-gel horizontal system (Hoefer) was used. DNA samples were mixed with a 2XHSE loading buffer, and RNA samples were mixed with an RNA-loading buffer. The prepared samples were loaded on 1-2% agarose gel and subjected to electrophoresis in 1x TBE buffer at a constant current of 50 mA. Before loading on the gel, the RNA samples were denatured at 85°C for 3 minutes and incubated on ice for 1 minute.

3.8 ChIP – Chromatin Immunoprecipitation

Plant tissue before Chromatin Immunoprecipitation experiments was crosslinked with formaldehyde. Briefly, the harvested tissue was placed in 50 ml tubes and kept on ice. 1x PBS with 1% formaldehyde was added to the kept tissue (~ 37.5 ml). A nylon mesh was placed in the tube to keep the tissue soaked in the buffer. To infiltrate the bugger into the plants tissue, a vacuum was applied twice for 10 minutes each. Next, a glycine solution was added to a final concentration of 0.125 M, and a vacuum was applied for another 5 min. After infiltration, the tissue was rinsed with water and dried with paper towels. Next, the tissue was ground in mortar with liquid nitrogen and stored at -80°C.

Nuclei isolation buffer I (40 ml) was added to 4 g of the grinded tissue. The tissue was re-suspended by vigorous vortexing and shaking. The sample was filtered into a new 50 ml tube using a Miracloth. The Miracloth was washed with 10 ml of Nuclei isolation buffer I. Next, the sample was centrifuged (15 min, 4,000 g, 4˚C).

The supernatant was discarded, and a nuclei pellet was re-suspended using 1 ml of cold Nuclei isolation buffer II. The sample was transferred to a 1.5 ml tube and centrifuged (5 min, 2,000 g, 4˚C). This step was repeated two more times (or more if the pellet was still green). The pellet was re-suspended using 300 µl of cold Nuclei isolation buffer II and layered on top of 900 ml of cold Nuclei isolation buffer III in a 1.5 ml tube. The sample was centrifuged (30 min, 16,000 g, 4˚C), and the supernatant was discarded.

The pellet was re-suspended in 300 µl of cold Sonic buffer. 10 µl of the sample was pipetted into a new tube as a control sample. The rest of the sample was sonicated using Diagenode Bioruptor (high mode, 30s/30s). 10 µl of the sample was pipetted into a new tube and treated with 1 µl of Proteinase K (15 min 65˚C) and 1 µl of RNase A (5 min, 37˚C). Both control samples (taken before and after sonication) were run in an agarose gel to check sonication efficiency. If the DNA band was visible as a 200-500 bp, a smear of the sample was taken for further analysis.

The sonicated chromatin was centrifuged (5 min, 3000 g, 4˚C). The supernatant was collected to a new "low binding" eppendorf tube and diluted with one volume of IP buffer. Sample concentration was measured using a Nanodrop. The difference in concentration between samples was equalized using a mixture (1:1) of Sonic and IP buffer. Next, 10% of the sample was pipetted into a new tube and frozen as an "input" sample. 50 µg of chromatin was used for one immunoprecipitation experiment. An appropriate antibody was added to the chromatin sample, and the mixture was incubated for 12-16 h at 4˚C, with rotation.

After overnight incubation, protein G magnetic beads (50 µl for one IP sample) were prepared by washing them three times with an IP buffer. Washed beads were added to the IP sample, and the mixture was incubated for 1 hour at 4˚C, with rotation. Immunoprecipitated chromatin was collected using a magnetic separator. The beads were washed for 5 min with cold buffers: two times with a Low-salt buffer, once with a High-salt buffer, once with a LiCl buffer, and twice with a TE buffer. After the last wash, the sample was transferred into a new tube, and the beads were collected using a magnetic separator.

Before qPCR analysis, DNA was eluted and de-crosslinked. Briefly, 300 µl of 10% Chelex was added to the sample, and the solution was incubated in a thermomixer (95˚C, 1,400 rpm, 10 min). 300 µl of 20% Chelex was added to a previously frozen input sample, and the sample was treated in the same manner as the IP sample. After the incubation

samples were centrifuged (5 min 14,000 g) and the supernatants were collected into new tubes. the IP sample was diluted 2x and the input sample 5x with water.

RESULTS

Post- Transcriptional regulation of biogenesis of microRNAs

1 Analysis of *MIR*163 gene expression in the *Arabidopsis* T-DNA insertion mutant

To learn about the influence of intron presence and/or splicing on microRNA biogenesis in plants, we needed an *Arabidopsis thaliana* mutant plant with disrupted expression of a *MIR* gene containing an intron. In addition, this *MIR* gene should be in one copy of the *Arabidopsis thaliana* genome. We have searched the publically available seed stock of SALK T-DNA insertion mutants (The European Arabidopsis Stock Center, www.arabidopsis.info). We focused on mutants where T-DNA insertion was found in a promoter region close to the transcription-initiation start of *MIR* genes. A SALK_0034556 mutant has T-DNA insertion in the promoter region of the *MIR163* gene. *MIR163* is a single-copy gene located on chromosome 1 between the At1g66720 and At1g66730 genes. *MIR163* contains one intron, and the microRNA is embedded in the first exon. Kurihara & Watanabe observed that two types of pri-miRNA163 are present in *Arabidopsis thaliana* (type 1 and type 2). Both have a polyadenylation signal, but the polyadenylation signal of type 1 is located in the intron of type 2, and type 2 pri-miRNA163 arises by splicing the intron having a polyA site in the last exon [Kurihara & Watanabe, 2004]. Schematic representation of the *MIR163* gene and two types of pri-miR163 are illustrated in Figure 6.

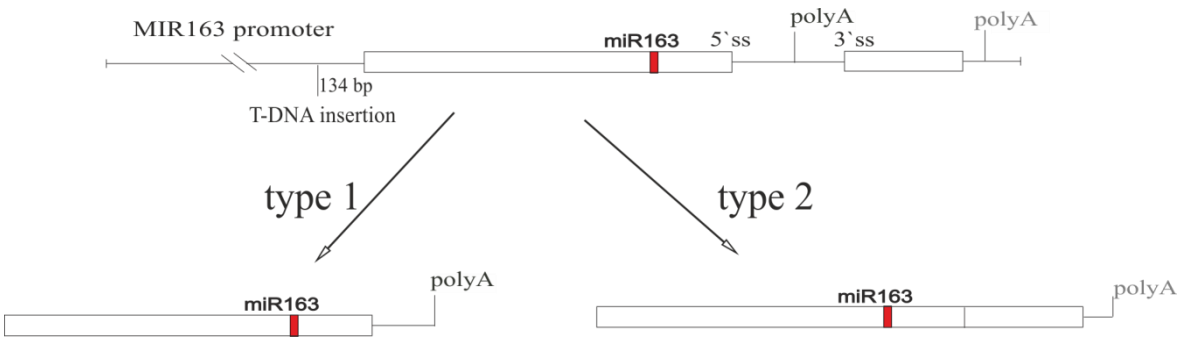


Figure 6 Schematic representation of the *MIR163* gene and two types of pri-miR163 observed by Kurihara & Watanabe in 2004. The T-DNA insertion, splicing sites, and position of poly(A) sites are shown.

To verify if the tested plants were indeed *mir163* null mutants, we studied the level of mature microRNA163 in SALK_0034556 as compared to wild-type plants. RNA for Northern blot hybridization was isolated from three different plant tissues (leaves, stems, and inflorescences) from the mutant plants. The results of the Northern blot experiment are presented in Figure 7. The analysis showed that, although the signal from miR163 is very strong in the T-DNA insertion mutant SALK_0034556 in wild-type plants, the level of mature miR163 is below detection by Northern blot hybridization.

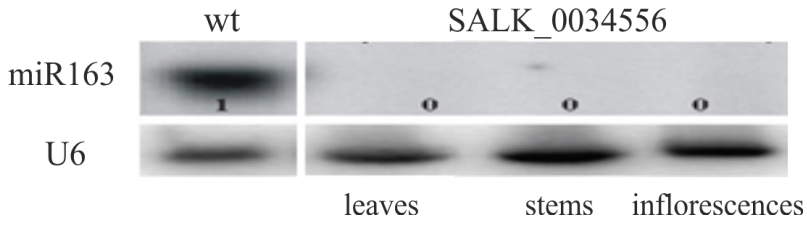


Figure 7 The level of mature microRNA163 (24 nt in length) recorded by Northern blot hybridization in different organs of the SALK_0034556 mutant and in wild-type plants. U6 snRNA served as a loading control.

We were also interested if T-DNA insertion would abolish the transcription from the *MIR163* gene. To verify the presence of pri-miRNA163 in the mutant's tissues, we performed a qPCR analysis. Reverse transcription was performed with an oligo(d)T primer, and the qPCR was conducted with primers amplifying a fragment of the first exon of the pri-miRNA163. As a consequence, both types of pri-miRNA163 were included in the analysis. RNA samples for qPCR analysis were the same as those used for detection of mature miR163 in the mutant and wild-type plants. The results of the qPCR experiment are presented in Figure 8. The analysis showed that the level of pri-miRNA163 in all analyzed mutant tissues is severely decreased in comparison to wild-type plants. When the level of pri-miRNA163 recorded in wild-type plants was arbitrarily set as a hundred percent, the relative expression level of pri-miRNA163 in SALK_0034556 reached only 0.34%. This means that the primary transcript of the *MIR163* gene is over 300 times less abundant in the mutant than in wild-type plants.

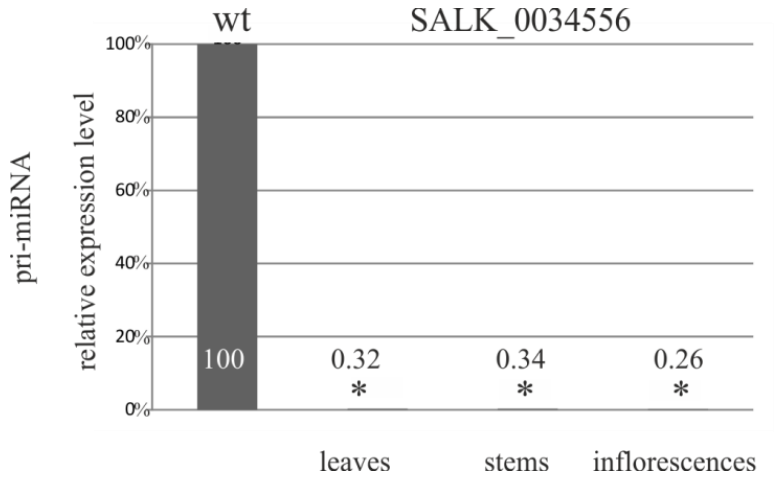


Figure 8 The level of pri-miRNA161 recorded by qPCR in different organs of the SALK_0034556 mutant and in wild-type plants. Asterisks indicate statistical significance difference between the indicated sample and control wild-type plants (Mann-Whitney test, $p<0.05$).

We also analyzed phenotypic traits of wild-type and mutant plants. We did not observe any phenotypic differences between mutant and wild-type plants, such as different plant stature, the shapes of rosette and stem leaves, time of flowering, or fertility (data not shown).

Altogether, the obtained results showed that the analyzed SALK_0034556 mutant plants were *mir163*-null mutants. In 2011, Danny Ng and his colleagues described a CS879797 mutant with T-DNA insertion in the *MIR*163 gene, resulting in a *mir163* null-mutant phenotype [Ng *et al.* 2011]. For this reason, we called our SALK_0034556 plant a *mir163-2* mutant.

2 Intron and splicing are required for proper miR163 biogenesis in *A. thaliana*

To evaluate the significance of introns in pri-miRNAs and their roles in the biogenesis of mature microRNAs, we constructed and introduced original or mutated *MIR163* gene variants under a native *MIR163* promoter in the *mir163-2* mutant background. Five constructs where generated:

1) IVSwt is a construct representing a native-gene version of *MIR163*.

2) ΔIVS is a construct representing an intronless version of the *MIR163* gene. The intronic fragment (154 bp) was removed from the construct using the PCR method (details described below).

3) IVSΔ5’ss is a construct representing the *MIR163* without a functional 5’ splice site. Six original nucleotides (three last nucleotides of the first exon and three first nucleotides of the intron) were changed to a tract of six thymidines. Recognition of 5’ss by U1 snRNP is the first step in the assembly of a functional spliceosome. This recognition is based on RNA:RNA hybridization between U1 snRNA and pre-mRNA/pri-miRNA 5`ss. Mutation in the IVSΔ5’ss construct should prevent spliceosome assembly.

4) IVSΔ3’ss is a construct representing *MIR163* without a functional 3’ splice site. Five original nucleotides (three last nucleotides of the intron and two first of the second exon) were changed to a tract of five thymidines. Mutation in the IVSΔ3’ss construct should prevent splicing of the intron.

5) IVSmut is a construct representing *MIR163* without either functional 5’ or 3’ splice sites. Mutations were prepared in the same manner as the IVSΔ5’ss and IVSΔ3’ss constructs.

Schematic representation of the constructs described above can be found in Figure 9. In all constructs, a native *MIR163* promoter was used. The promoter region represents a 1,400-base-pair-long fragment. A genomic fragment between the transcription initiation site of the *MIR163* gene and the 3’ end of the closest annotated gene was selected as a whole promoter region.

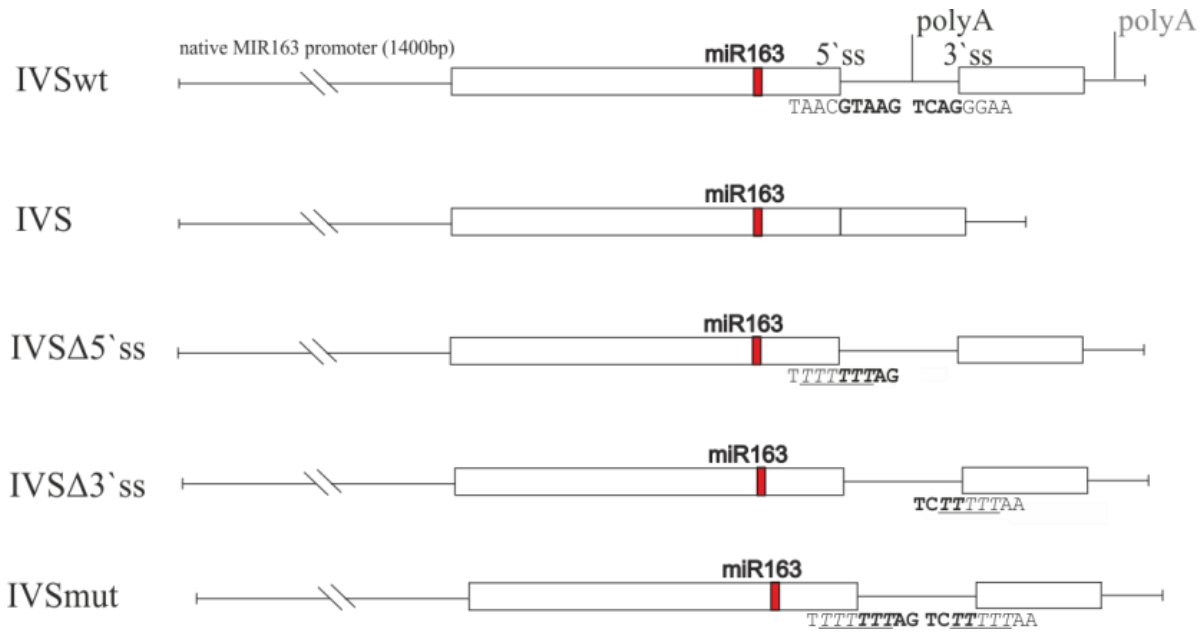


Figure 9 Schematic representation of the *MIR163* gene variants. Splicing sites and positions of proximal and distal poly(A) sites are marked. Nucleotide sequences of wild-type and mutated splice sites are shown.

The construct representing a native gene of *MIR163* (IVSwt) was amplified by a PCR using A01 and A02 primers and genomic DNA from wild-type plants as a template (for primers, see the table from Chapter 3 in Materials). The obtained PCR product was cloned into a pENTR/D – TOPO plasmid. To create an intron-less version of the *MIR163* gene (ΔIVS), three separate PCR reactions were conducted. In the first reaction, the promoter and first exon of the *MIR*163 gene were amplified using A03 and A04 primers. The reverse primer (A04) for this reaction was designed in such manner that the sequence of its 5` end was complementary to the first 19 nucleotides of the second exon of the *MIR163* gene. In the next PCR reaction, the second exon was amplified using A05 and A06 primers. The forward primer for this reaction (A05) was designed in such manner that its 5` end was complementary to the last 16 nucleotides of the first exon of the *MIR163* gene. PCR products were separated using agarose gel electrophoresis, excised from agarose gel, and used for the third PCR as templates. In this PCR, the forward primer from the first reaction (A03) and the reverse primer from second PCR reaction (A06) were used. The obtained final PCR product was separated using an agarose electrophoresis, excised from the gel and cloned into a pENTR/D-TOPO plasmid using NotI and AscI restriction sites. The strategy used during this construct preparation is presented in Figure 10. To

create a construct where the 5` splice site (IVSΔ5`ss) or 3` splice site (IVSΔ3`ss) is mutated, directed mutagenesis was conducted according to the manufacturer's protocol (primers listed in Material & Methods, Chapter 3).

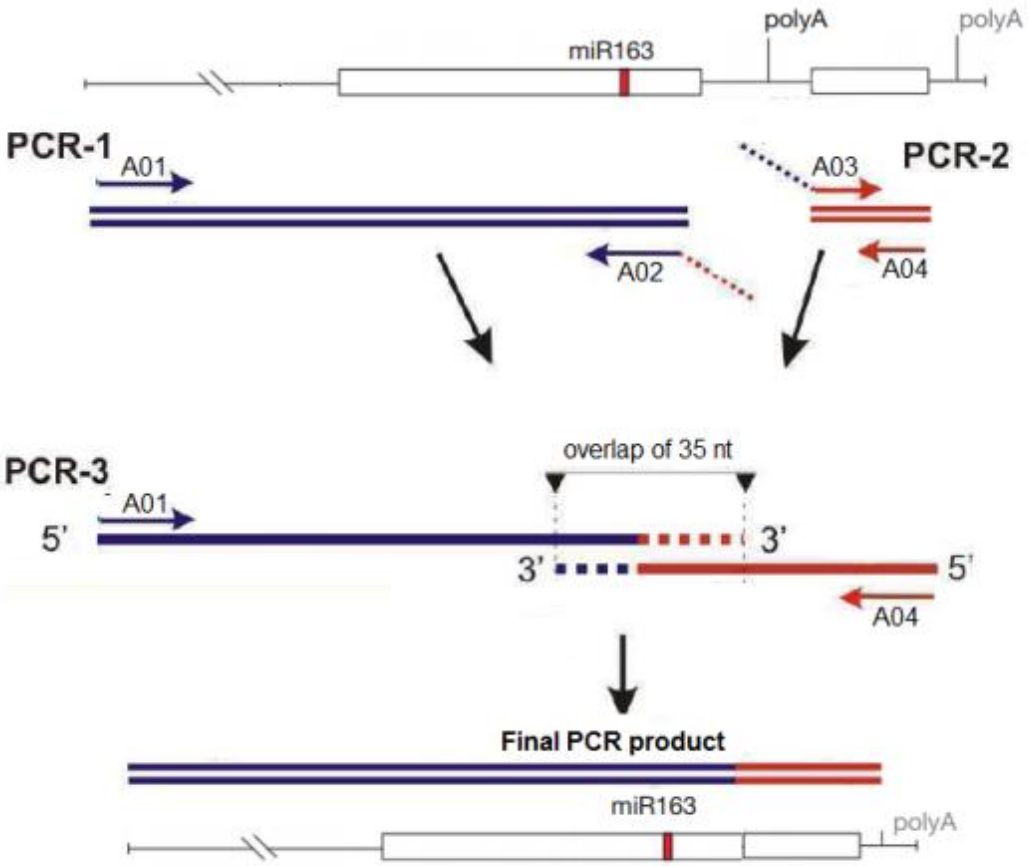


Figure 10 Schematic overview of the described procedure in the generation of the intron-less version of the *MIR163* gene.

The obtained IVSwt, ΔIVS, and IVSmut constructs were cloned to a pMDC99, and the IVSΔ5'ss and IVSΔ3'ss constructs were cloned to a pMDC123 binary vector using Clonase system from Life Technologies, followed by electroporation of the *Agrobacterium tumefaciens* AGL-1 strain. Binary vectors pMDC99 and pMDC123 differ only in a fragment that encodes a resistance gene that is used as a selection marker in transgenic plants. Plants transformed with a pMDC99 vector are resistant to the antibiotic hygromycin, and plants transformed with a pMDC123 vector are resistant to the herbicide Basta. Prepared strains of AGL-1 bacteria were used for transformation of the *mir163-2* mutant by the floral-dip method [Cloguh & Bent 1998].

After transformation, the plants were grown in a long-day photoperiod condition in a Sanyo chamber until the siliques turned yellow and then brown. The collected seeds (the

T1 generation of plants transformed with IVSwt, ΔIVS, or IVSmut constructs) were sieved on a ½-MS medium supplemented with hygromycin. The seeds were incubated in Petri dishes in a Sanyo chamber for three weeks. Next, green seedlings were transferred to soil and grown until the siliques matured. In the case of transformed plants carrying IVSΔ5'ss or IVSΔ3'ss constructs, the collected seeds were sown directly into the soil. After germination, the seeds were sprayed five times at two-day intervals with a 300 μM Basta solution. Arbitrarily selected plants from each T1 generation were grown until the siliques matured. During the growth of the plants, genomic DNA was isolated from one leaf of each selected T1 plant. Samples after DNA isolation were used to verify the presence of a transgene by PCR. Results of the PCR experiments are presented in Figure 11. Primers complementary to *hyg* or *bar* genes were used. The results of the PCRs confirmed the presence of transgenes in all of the selected plant lines.

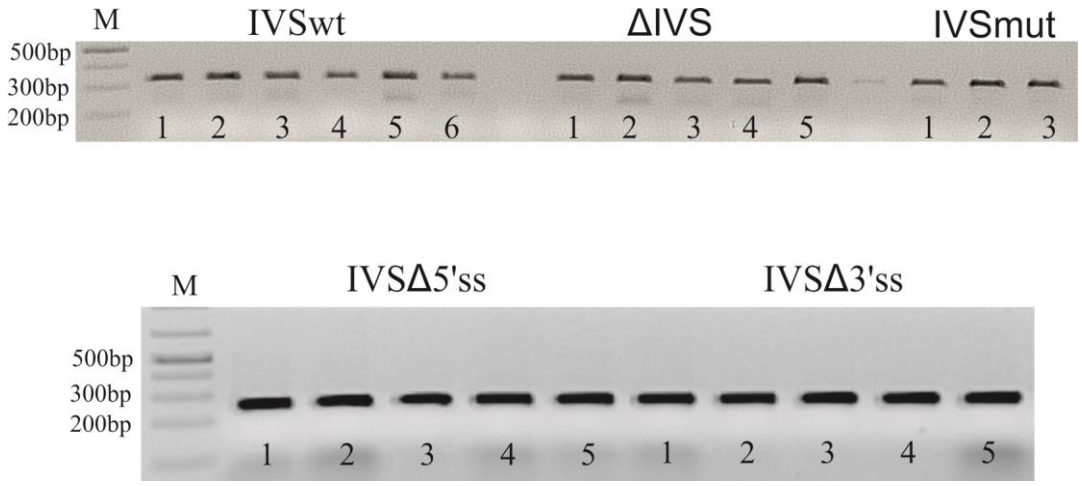


Figure 11 Agarose gel electrophoresis of PCR products confirming the presence of a transgene in selected plant lines. Numbers indicate independent transgenic lines, M – DNA ladder

All of the selected plants in the T1 generation should have only one copy of a transgene. To verify the T-DNA copy number, collected seeds from T1 generation were sown on a ½ MS medium supplemented with hygromycin or Basta. If plants were hemizygous, one-fourth of the plants grown in Petri dishes should not survive the selection. The ratio of green *versus* pale white seedlings was calculated, and in all analyzed cases, segregation confirmed that the plants from the T1 generation were indeed hemizygous. Next, green seedlings were transferred to soil and grown until the siliques matured. Seeds were collected individually from each plant. Among batches of the

collected seeds are hemizygous and homozygous plants. To learn which batch represents which genotype, seeds were again sawn on a ½-MS selected medium, and the ratio of green *versus* pale white seedlings was calculated. If the batch of seeds contained homozygous transgenic plants, all seedlings on a Petri dish should be viable. For each construct, at least two homozygous T1 lines of transgenic plants were selected.

During selection and propagation of the transgenic plants, phenotypic traits (plant stature, shape of rosette, and stem leaves) were observed. We did not observe any phenotypic differences between the transgenic lines, *miR163-2* mutant, or wild-type plants (data not shown).

Wild-type, *mir163-2* mutant, and the obtained transgenic homozygous seeds containing the IVSwt and ΔIVS constructs were sown on a ½-MS medium and grown in a Sanyo chamber for two weeks. RNA was isolated from whole plants using a TRIzol reagent. Next, after DNase I treatment, a semi-quantitive RT-PCR was conducted. Results of the semi-quantitative RT-PCR are shown in Figure 12. Analysis of the transgenic plants carrying the IVSwt construct (a wild type of the *MIR163* gene variant) showed that, in two independent transgenic lines, two isoforms of pri-miRNA163 can be detected. The same two isoforms of pri-miR163 were detected in wild-type plants. The isoforms correspond to unspliced and spliced pri-miRNA163. In the two ΔIVS independent homozygous transgenic lines expressing an intron-less *MIR*163, only one isoform of pri-miR163 was detected. The presence of the isoform that corresponds to the spliced transcripts of pri-miR163 confirmed that, after removal of the intronic sequence from the *MIR163* gene, no cryptic spliced sites were activated.

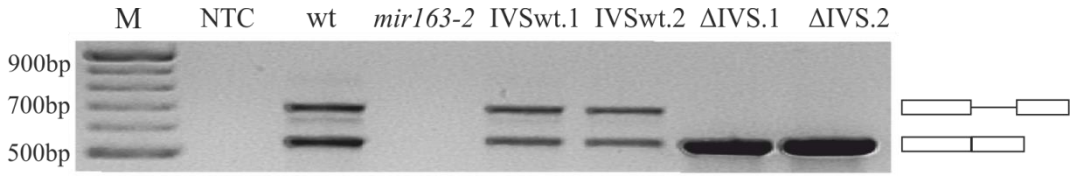


Figure 12 The results of the semi-quantitative RT-PCR analysis of wild-type, *mir163-2*, and transgenic lines carrying the IVSwt or ΔIVS constructs. Products of PCR were separated using agarose electrophoresis. M –DNA ladder, NTC- non template control.

We decided to quantify and compare pri-miR163 levels in the different lines of *Arabidopsis thaliana* transgenic plants. The cDNA samples that were used in the semi-quantitative RT-PCR analysis were also used for real-time PCR analysis. Primers for

amplification of pri-miR163 were complementary to the first exon of *MIR163*, and the level of transcripts of *ACT2* was used as a calibrator between samples in the qPCR analysis. Results of the qPCR analysis are shown in Figure 13. The obtained results showed that, in transgenic plants carrying the IVSwt construct, the level of pri-miRNA163 is the same as the level of pri-miR163 recorded in wild-type plants. However, the level of pri-miR163 in the transgenic plants carrying the ΔIVS construct was significantly higher than that of the wild-type plants.

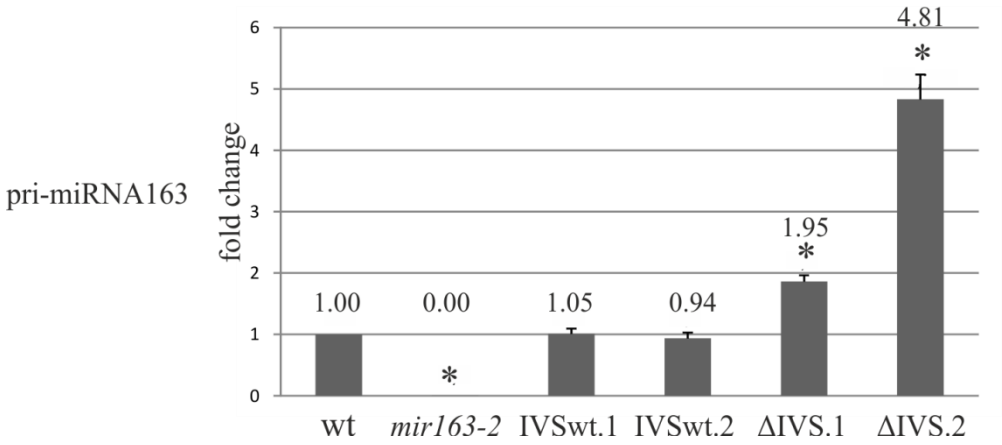


Figure 13 The results of the qPCR analysis of the wild-type, *mir163-2*, and transgenic lines carrying the IVSwt or ΔIVS constructs. Asterisks indicate statistical significance difference between the indicated sample and control wild-type plants (Mann-Whitney test, $p<0.05$).

One of the possible explanations for the elevated levels of pri-miR163 in the ΔIVS transgenic lines can be that the transcripts that rise from the intron-less variant of the *MIR163* gene are not efficiently processed by microRNA machinery complexes. If indeed the transcripts accumulate due to the fact that they are not efficiently processed to mature microRNA163, the level of 24-nt microRNAs should be downregulated in the ΔIVS transgenic plants as compared to wild-type or IVSwt plants. To verify this hypothesis, Northern-blot hybridization was performed. Total RNA was isolated from two-week-old seedlings and separated on a PAGE in the presence of 8M urea. In hybridization, a radioactive-labeled oligonucleotide probe was used that is complementary to miR163. As a marker for RNA-loading control, we tested the level of U6 snRNA using the appropriate radioactive-oligonucleotide probe (see table from Chapter 3 in Materials). The level of mature microRNA recorded in wild-type, *mir163-2*, and the transgenic lines is shown in

Figure 14. The level of microRNA detected in the wild-type plants was arbitrarily set as one hundred percent. The obtained results showed that the level of mature microRNA in the IVSwt transgenic plants and wild-type plants was similar. However, the level of microRNA163 in the transgenic plants carrying an intron-less version of the *MIR163* gene was almost three times lower than that of the wild-type or IVSwt plants.

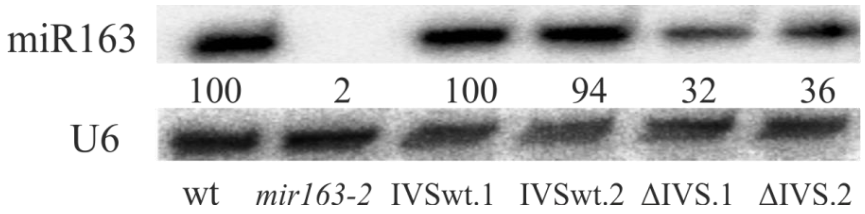


Figure 14 Results of Northern-blot analysis of the wild-type, *mir163-2*, and transgenic lines carrying the IVSwt or ΔIVS constructs. U6 snRNA served as a loading control.

The obtained results suggest that the intron of *MIR163* is required for the proper accumulation of miR163, but it is unclear whether this is caused by splicing or by an unknown stimulatory feature of a sequence motif within the intron. To answer this, we analyzed transgenic plants in which different versions of the *MIR163* gene with mutated splice sites were introduced in the *mir163-2* mutant background.

Wild-type and the obtained transgenic homozygous seeds containing the IVSΔ5'ss, IVSΔ3'ss, and IVSmut constructs were sown in soil and grown in a Sanyo chamber for three weeks. RNA was isolated from each whole plant using a TRIzol reagent. Next, after DNase I treatment, a semi-quantitive RT-PCR was conducted. In all transgenic lines, only one isoform of pri-miRNA163 was detected. The detected isoform corresponds to the unspliced transcript of the *MIR*163 gene. The analysis confirmed that splicing is inhibited after the inactivation of the 5'ss (IVSΔ5'ss) and 3'ss (IVSΔ3'ss) as well as after the mutation of both splice sites (IVSmut) in *MIR163*. Results of the semi-quantitative RT-PCR are shown in Figure 15.

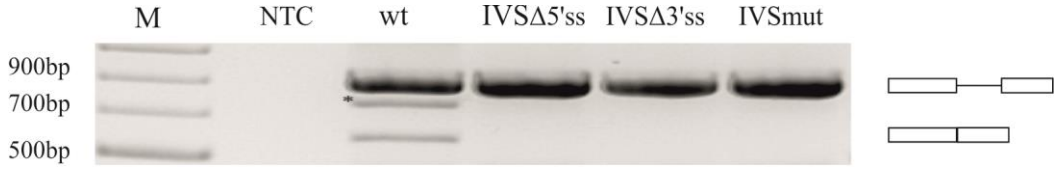


Figure 15 The results of the semi-quantitative RT-PCR analysis of the wild-type and transgenic lines carrying the IVSΔ5'ss, IVSΔ3'ss, or IVSmut constructs. Products of the PCR reactions were separated using agarose electrophoresis. M –DNA ladder, NTC- non template control. The asterisk points to an unidentified product.

To quantify and compare the pri-miR163 levels in the different lines of *Arabidopsis thaliana*, the prepared cDNA samples that were used in the semi-quantitative RT-PCR analysis were also used for real-time PCR analysis. Results of the RT- qPCR analysis are shown in Figure 16. The obtained results showed that, in the transgenic plants carrying the IVSΔ5'ss construct, the level of pri-miRNA163 is the same as the level of pri-miR163 recorded in wild-type plants. The level of pri-miR163 in transgenic plants carrying the IVSΔ3'ss construct was slightly (yet significantly) higher than that recorded in wild-type plants, and the level of pri-miR163 in transgenic plants carrying the IVSmut construct was severely decreased as compared to the wild-type plants.

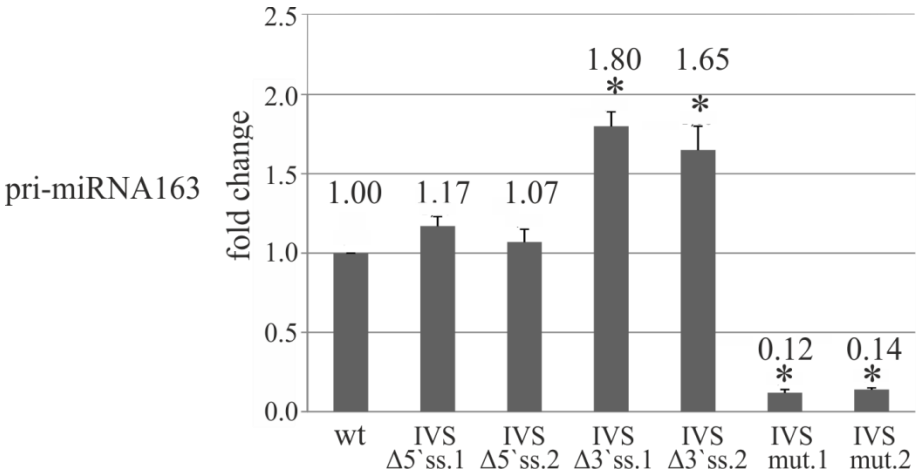


Figure 16 The results of the qPCR analysis of wild-type and transgenic lines carrying the IVSΔ5'ss, IVSΔ3'ss, or IVSmut constructs. Asterisks indicate significant statistical difference between the indicated sample and control wild-type plants (Mann-Whitney test, $p<0.05$).

To analyze the level of mature miR163 in the wild-type and transgenic plants, Northern-blot hybridization was performed. Results of the analysis are shown in Figure 17.

The obtained results showed that the level of mature microRNA in the IVSΔ5'ss transgenic plants is lower than that recorded in wild-type plants. However, the level of microRNA163 in the transgenic plants carrying the mutation in the 3'ss of the *MIR163* gene was almost the same as the level of micro163 in the wild-type plants. Mutation of both splice sites of the *MIR163* gene almost completely abolished the production of miR163. The level of miR163 in two independent transgenic lines reached only 16-17 percent of the level of microRNA163 recorded in wild-type plants.

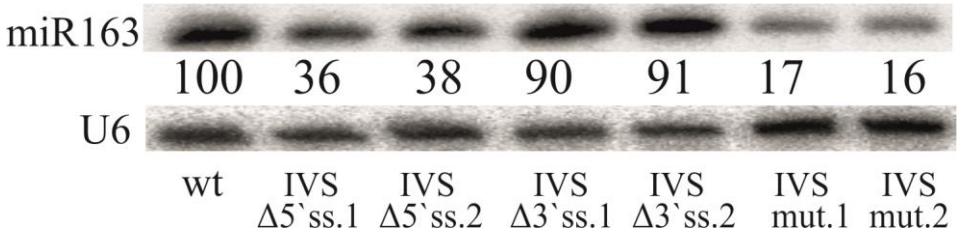


Figure 17 The results of the Northern-blot analysis of the wild-type and transgenic lines carrying the IVSΔ5'ss, IVSΔ3'ss, and IVSmut constructs. U6 snRNA served as a loading control.

The obtained results suggest that the splicing of *MIR163*'s intron is required for the accumulation of the proper level of miR163. Additionally, the results showed that the accumulation of miR163 is far less affected by mutation in the 3` splice site than in the 5` site. Moreover, mutation of both splice sites had an additive effect on decreasing the accumulation of miR163, correlating with strong downregulation of the mutant pri-miRNA163 level.

3 Intron and splicing are required for proper biogenesis of another *MIR* gene, *MIR161*.

To evaluate whether the intron/splicing stimulatory effect on the biogenesis of plant miRNAs is not specific to the *MIR163* gene, we tested another *MIR* gene. We decided to use *Nicotiana benthamiana* plants and perform a transient expression on a selected *MIR* gene. *MIR161* in *Arabidopsis thaliana* is a single-copy gene containing one intron, and microRNA161 is embedded in the first exon. Analogous constructs like in the case of *MIR163* were prepared. The main difference between the constructs is that, in the case of *MIR161*, we used a strong 35S promoter from a CaMV that is included in a pMDC32 binary vector in all constructs. Schematic representation of the prepared *MIR161* gene constructs can be found in Figure 18.

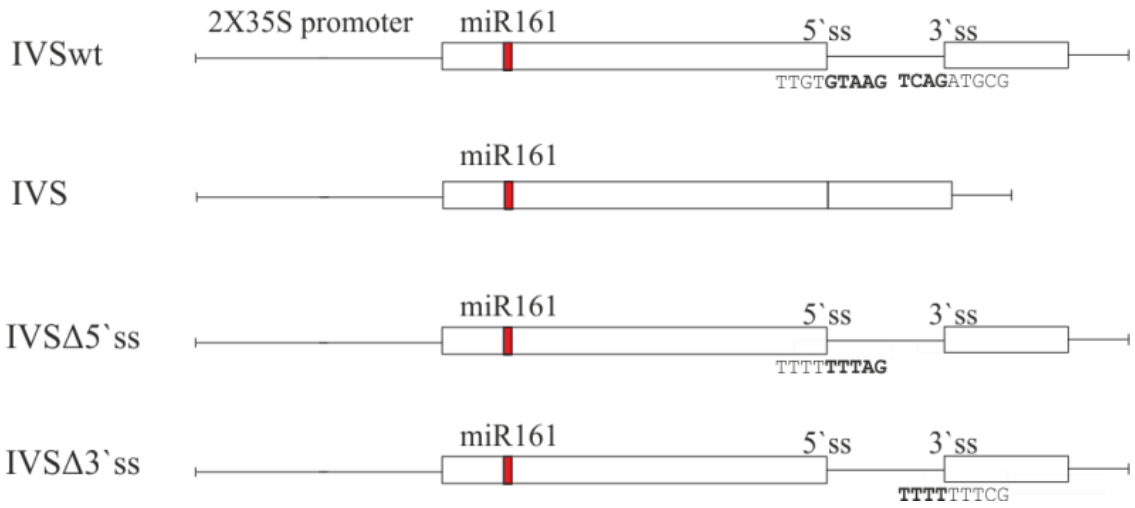


Figure 18 Schematic representation of *MIR161* gene variants. The sequences of wild-type and mutated splice sites are shown.

Leaves of five-week-old *Nicotiana* plants where infiltrated with an *Agrobacterium* AGL1 strain carrying one of the prepared variants of the *MIR161* gene. Next, 72 hours after infiltration, the leaves were ground in liquid nitrogen, and total RNA was isolated from the collected samples. After DNase I treatment, a semi-quantitive RT-PCR was conducted. Results of the semi-quantitative RT-PCR are shown in Figure 19. In the analyzed samples, only one isoform of pri-miR161 was detected. In leaves expressing miR161 from IVSwt, IVSΔ5'ss, or IVSΔ3'ss constructs, the recorded isoform corresponded to the unspliced isoform of pri-miR161, whereas in leaves expressing miR161 from the ΔIVS construct, the recorded isoform corresponded to a spliced isoform of pri-miR161. The obtained results confirmed that, after removal of the intronic sequence from the *MIR161* gene, no cryptic splicing sites were activated.

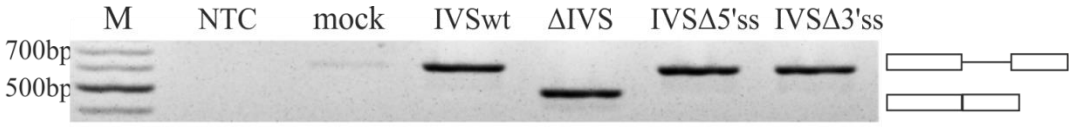


Figure 19 The results of semi-quantitative RT-PCR analysis of samples from leaves infiltrated with *Agrobacterium* AGL1 carrying one of the prepared variants of the *MIR161* gene. Products of PCRs were separated using agarose electrophoresis. Mock – samples from the leaves infiltrated only with buffer, NTC- non-template control, M –DNA ladder.

However, it was not clear why the spliced isoform of the pri-miRNA161 was not detected in samples collected from the leaves transformed with the IVSwt construct. We decided to perform a semi-quantitive RT-PCR analysis on a pri-miR161 in *Arabidopsis thaliana* wild-type plants and in microRNA biogenesis mutants. The results of the semi-quantitative RT-PCR are shown in Figure 20. The obtained results showed that the isoform that corresponded to the spliced isoform of pri-miR161 was detected only in samples from *hyl1-2* mutants. The analysis also showed that pri-miR161 overaccumulates in *hyl1-2*, *se-1*, *cbp20*, *cbp80*, and *cbc* mutants. In 2009, Bogna Szarzyńska and her colleagues analyzed the ratio between spliced and unspliced isoforms of pri-miR161. They carried out an RT-PCR using fluorescently labeled primers and analyzed products by capillary electrophoresis. Their analysis showed that there is only a slight difference in the ratio of unspliced/spliced pri-miRNA between wild-type and *hyl1* plants. Szarzyńska's (and our) results suggest that a steady-state level of spliced isoforms of pri-miR161 in wild-type plants is far less abundant than in unspliced isoforms. These results can explain why no spliced isoform was detected by the semi-quantitive RT-PCR in *Nicotiana* leaves that express the IVSwt version of the MIR161 gene.

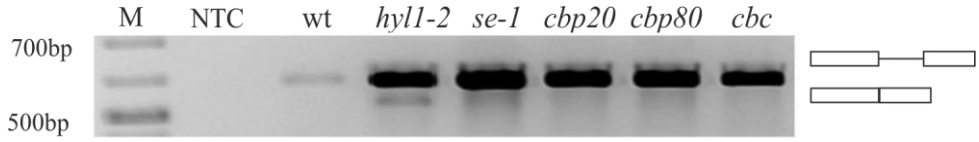


Figure 20 The results of the semi-quantitative RT-PCR analysis of pri-miRNA161 in wild-type and microRNA biogenesis mutants. Products of the RT-PCR were separated using agarose electrophorese. M –DNA ladder, NTC- non-template control.

We decided to quantify and compare a steady state level of pri-miR161 that is expressed from different variants of *MIR161*. The cDNA from *Nicotiana* samples were used in real-time PCR analysis. Results of the qPCR analysis are shown in Figure 21. The obtained results showed that the level of pri-miR161 expressed from the IVSwt construct was the same as in leaves transformed with IVSΔ5`ss or IVSΔ3`ss constructs. In the ΔIVS-transformed leaves, the level of pri-miR161 was significantly higher than in other samples. The expression of the hygromycin phosphotransferase gene (*hptII*) served as a positive control of agroinfiltration and as a reference gene in qPCR calibration.

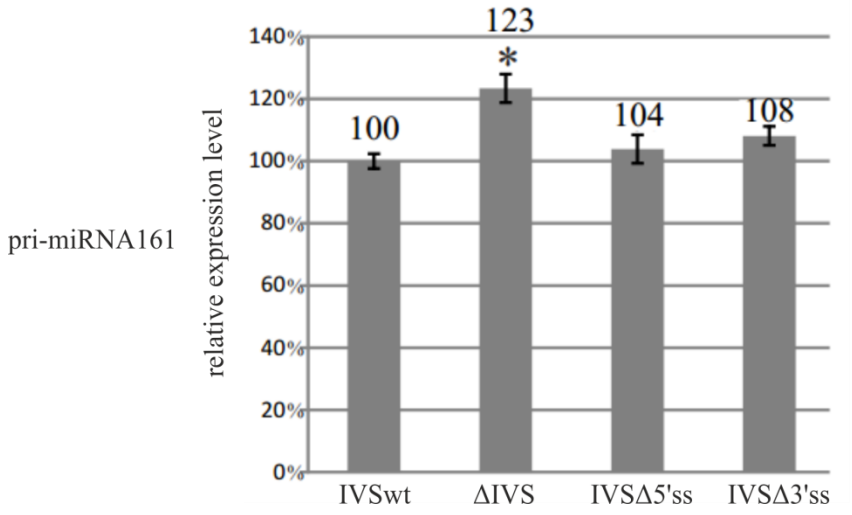


Figure 21 The results of qPCR analysis of RNA isolated from the leaves of *Nicotiana* infiltrated with *A. tumefaciens* carrying the IVSwt, ΔIVS, IVSΔ5'ss, or IVSΔ3'ss constructs. Asterisks indicate significant statistical difference between the indicated sample and control wild-type plants (Mann-Whitney test, $p<0.05$).

To analyze the level of mature miR161 in the collected *Nicotiana* RNA samples, Northern-blot hybridization was performed. The results of the analysis are shown is Figure 22. The analysis of mature miRNA by the Northern blot showed that the level of miRNA161 is severely decreased in samples where expression was driven from an intron-less version of the gene, (the ΔIVS gene variant). The level of miR161 was also dramatically decreased when the 5` splice site was mutated (the IVSΔ5'ss gene variant). However, when the 3` splice site was mutated, only a mild effect on microRNA accumulation was observed.

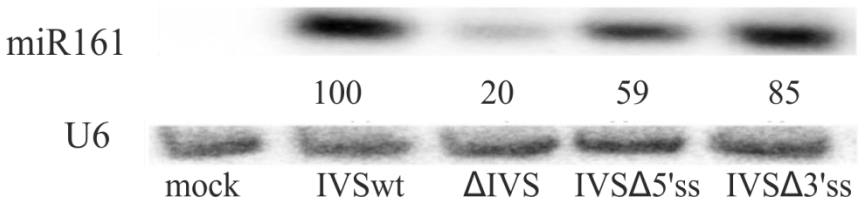


Figure 22 The results of Northern-blot analysis of RNA isolated from the leaves of *Nicotiana* infiltrated with *A. tumefaciens* carrying the IVSwt, ΔIVS, IVSΔ5'ss, or IVSΔ3'ss constructs. Mock – samples from leaves infiltrated only with a buffer. U6 snRNA served as a loading control.

The obtained results from the transient expression experiment of *MIR161* are similar to those observed in the case of *MIR163*. Removal of the intronic sequence from the *MIR163* gene lead to three times the decreased level of mature microRNA, and removal of the intronic sequence from *MIR161* lead to a decreased level of mature microRNA (down to 20%, compared to the level of microRNA rising from native variants of *MIR* genes). In both cases (*MIR*163 and *MIR*161), mutation of the 3' splice site abolished the accumulation of mature microRNA, but this effect was more prominent when the 5' splice site was mutated (please compare Figures 12 and 17). Altogether, these results suggest a general stimulatory effect of introns on the biogenesis of plant miRNAs.

4 Two poly(A) sites are used to terminate transcription of the *MIR163* gene

The presence of two polyadenylation signals (PASs) in pri-miR163 transcripts prompted us to investigate an additional function of U1 snRNP (independent of its function in splicing). This additional function of U1 snRNP is a protection of transcripts from premature cleavage and polyadenylation (PCPA) [Kaida *et al.* 2010]. Kaida and his colleagues showed that cryptic polyadenylation signals were activated in HeLa cells when the binding between U1 snRNA and 5' splice sites were blocked. This cryptic PAS typically is within an intron, and their activation results in the failure in production of full-length pre-mRNAs from the majority of genes.

If the U1 snRNP protects pri-miR163 transcripts from premature cleavage and polyadenylation in transgenic plants where an IVSΔ5'ss gene variant was introduced, the ratio between type 1 (proximal) and type 2 (distal) of pri-miR163 transcripts should be changed as compared to the ratio recorded in wild-type plants (see Figure 5). We used a real-time PCR method to calculate the levels of type 1 and type 2 (unspliced and spliced) transcripts and to compare the ratio of polyA site selection in all variants of *MIR163* gene transcripts ended with a poly(A) tail. As expected, the calculated ratio in transgenic plants expressing pri-mi163 from the intron-less (ΔIVS) version of *MIR163* demonstrates that only the distal poly(A) site was used. The proximal poly(A) site was removed along with the intronic sequence, so this result can be used as proof of the principle of our method in calculating the ratio between type 1 and type 2 isoforms. The obtained qPCR results for wild-type plants and transgenic plants expressing IVSΔ5'ss, IVSΔ3'ss, and IVSmut variants of the *MIR163* gene are presented in Figure 23. The qPCR results showed that, in the wild-type plants, the proximal poly(A) site is used for roughly 40% of all *MIR163* transcripts. The same frequency of proximal poly(A) selection was observed in the case of

transgenic plants carrying the IVSΔ3'ss *MIR*163 mutant gene. However, in the case of transgenic plants carrying the IVSΔ5'ss mutant version of the *MIR*163 gene, the proximal poly(A) site was used in about 82-84% of all *MIR163* transcripts. Additionally, in the transgenic plants where pri-miR163 is driven from the gene variant where both splice sites are mutated (IVSmut), 95-96% of all pri-miR163 transcripts were terminated at the proximal poly(A) site. Therefore, similar to human cells, the binding of the 5`ss of pri-miR163 by a U1 snRNP might inhibit the usage of the proximal poly(A) site.

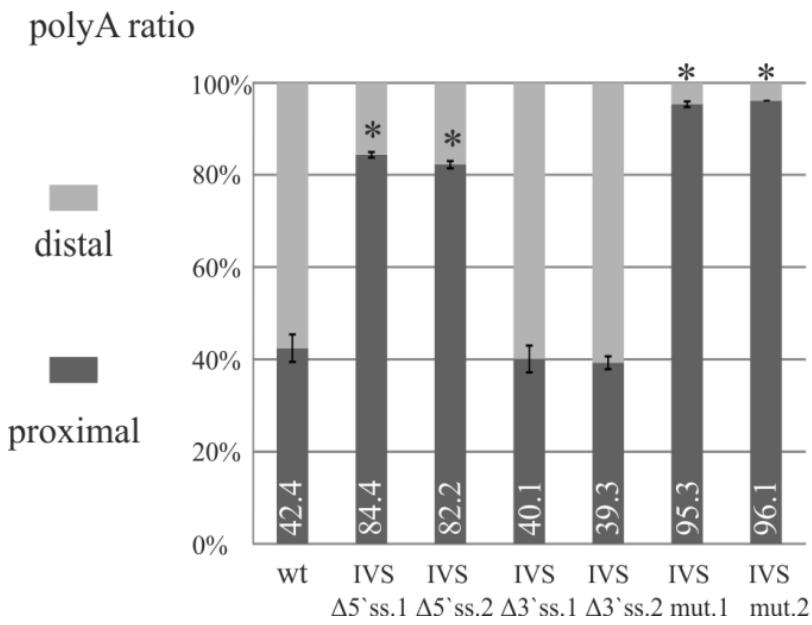


Figure 23 The ratio of proximal versus distal poly(A) site usage determined in wild-type plants and *MIR*163 variant lines. Asterisks indicate a significant difference between the indicated and IVSwt samples (Mann-Whitney test, $p<0.05$).

5 SR proteins affect miRNA biogenesis

Plant serine/arginine-rich (SR) proteins are important splicing factors that act as positive regulators in splicing. These proteins are evolutionarily highly conserved, and are characterized by the presence of one or two RNA-binding domains and an arginine/serine-rich (RS) domain. It was described that base pairing of U1 snRNA with a 5' splice site is stimulated by a member of the SR protein family - SF2/ASF [Jamison *et al.* 1995].

We decided to test different sr-null mutants of *Arabidopsis thaliana* and verify the hypothesis that splicing the pri-miRNA163 is an important step in miR163 biogenesis. In our experiments, we also included mutants known to be involved in miRNA biogenesis and mRNA splicing. Firstly, a semi-quantitive RT-PCR analysis of pri-miR163 in the described mutants was performed. The obtained results are presented in Figure 24. In the *hyl1-2* mutants, we observed that the spliced isoform accumulates as compared to wild-type plants. In the *se-1*, *cbp20*, *cbp80,* and *cbc* mutants, an unspliced isoform of pri-miR163 accumulates as compared to wild-type plants, and in the sr-null mutants, the level of the two splicing isoforms is more or less the same as in wild-type plants.

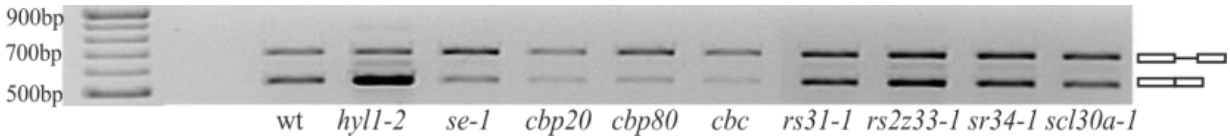


Figure 24 The results of semi-quantitative RT-PCR analysis of pri-miRN163 in wild-type plants, microRNA biogenesis mutants, and sr-null mutants. Products of RT- PCRs were separated using agarose electrophoresis.

To precisely quantify the ratio of unpliced versus spliced isoforms of pri-miR163, we performed qPCR in three biological replicates. The obtained results of qPCR are presented in Figure 25. The analysis confirmed the previous results from the RT-PCR experiment that spliced isoforms accumulate in the *hyl1-2* mutant, and in the *se-1*, *cbp20*, *cbp80*, and *cbc* mutants, unspliced isoforms accumulate as compared to wild-type plants. However, after comparing the sr-null mutants with the wild-type plants, we observed a slightly (yet statistically significant) decreased level of the spliced isoform of the pri-miR163.

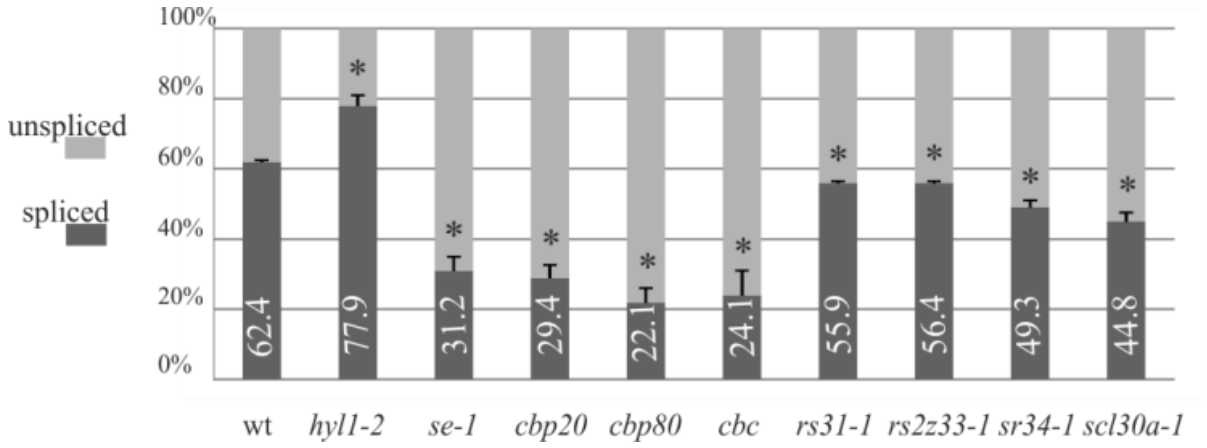


Figure 25 Ratio of spliced versus unspliced isoform abundance determined in wild-type plants, microRNA biogenesis mutants, and sr-null mutants. Asterisks indicate a significant difference between the indicated and IVSwt samples (Mann-Whitney test, $p<0.05$).

To learn if the ratio of spliced versus unspliced isoforms of the pri-miR163 has an influence on the level of mature microRNA163, we performed Northern-blot hybridization. The obtained results are presented in Figure 26. In the mutants known to be involved in miRNA biogenesis (*hyl1-2*, *se-1*, *cbp20*, *cbp80*, and *cbc*) the level of mature microRNA163 was severely decreased as compared to wild-type plants (as expected). Additionally, the level of mature microRNA decreased to 62 percent in sr-null mutants. These observations support our conclusion that splicing stimulates miRNA production from intron-containing pri-miRNA163.

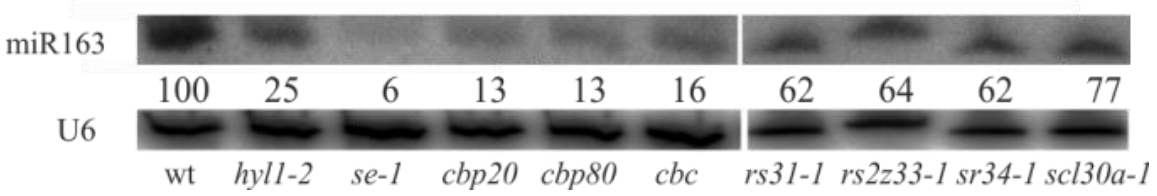


Figure 26 The results of Northern-blot analysis of mature miR163 in wild-type plants, microRNA biogenesis mutants, and sr-null mutants. U6 snRNA served as a loading control.

Moreover, we wondered if we could repeat the same results in the case of another intron-containing pri-miRNA. We performed Northern-blot hybridizations for the detection of mature microRNA161 and 171 in the described mutants (the genes encoding

both miRNAs contain one or more introns). MicroRNA161 can arise only from a single locus, and its pri-miRNA contains one intron. However, microRNA171 can arise from three different *loci* (called *MIR*171A, *MIR*171B and *MIR*171C); additionally, the transcribed pri-microRNAs from all three *loci* contain introns. The obtained results are presented in Figure 27. The analysis showed that the levels of microRNA161 and 171 are decreased in all of the analyzed mutants, which again strengthens the hypothesis that splicing stimulates miRNA production from intergenic intron-containing *MIR* genes.

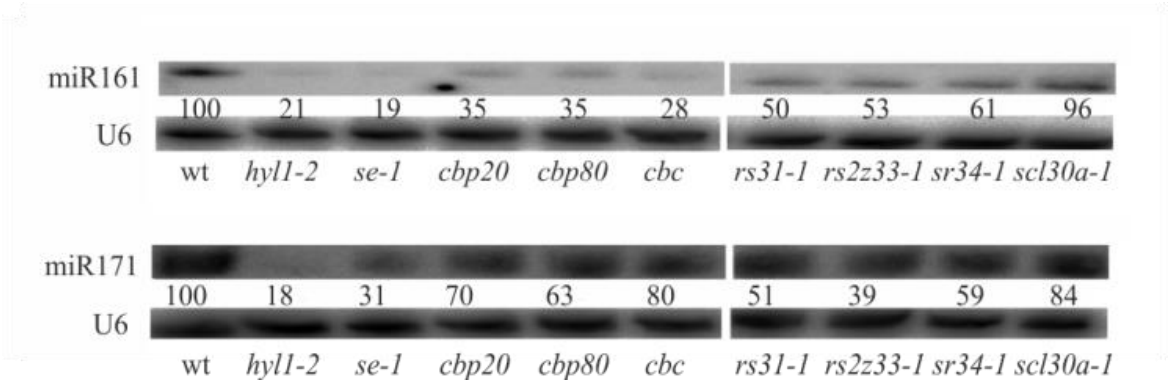


Figure 27 Northern-blot analysis of mature miR161 or miR171 in wild-type plants, microRNA biogenesis mutants, and sr-null mutants. U6 snRNA served as a loading control.

6 miR163 targets a SAM-dependent methyltransferase

MicroRNAs work as fundamental negative sequence-specific regulators of gene expression. To understand the function of a microRNA, it is essential to predict and validate microRNA targets. According to the ASRP database (*Arabidopsis thaliana* Small RNA Project), a transcript of the S-adenosyl-L-methionine-dependent methyltransferase (At1g66690) gene can be a target of microRNA163. It is very important to experimentally confirm that predicted *in silico* mRNA is the real target of any given microRNA. In our studies, we decided to perform RNA ligase mediated – 5` rapid identification of cDNA ends (5` RLM-RACE). In this approach, if the At1g66690 is cleaved by an miR163-guided mechanism, we should be able to find a 3` degradation intermediate of a cleaved mRNA transcript from the *AT1G66690* gene. A specific primer complementary to the At1g66690 transcript was used in the first and second rounds of amplification. After the first step of amplification, the RT-PCR mixture was subjected to agarose gel electrophoresis; however, no products were visible. For the second step of amplification, a 0.1 volume of the first-

round reaction mixture was used as a template. Half of the volume of the reaction mixture was subjected to agarose gel electrophoresis, and the rest was used for cloning the obtained products into a pGEM T-Easy vector. In the 5`RLM-RACE experiments, we also included *mir163-2* and *xrn4-3* mutants. The reason to use the *xrn4-3* plants was that 3` cleavage fragments of selected miRNA targets often accumulate in this mutant. The images of the agarose gel electrophoresis are shown in Figure 28.

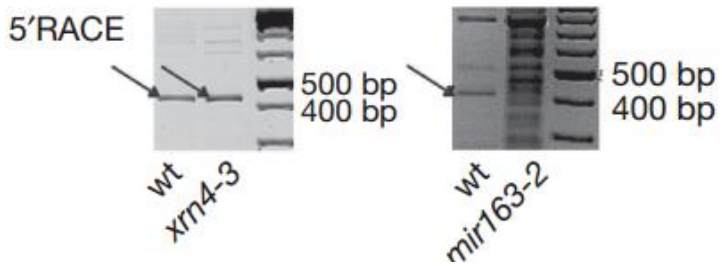


Figure 28 The agarose gel electrophoresis of the second amplification step of the 5'RLM-RACE of the AT*1G66690* mRNA slicing products. The PCR products at the expected size are marked with arrows.

After cloning the obtained products from the wild-type plants, eight plasmids were sequenced and analyzed. The obtained results after sequence analysis of the plasmids are shown in Figure 29. Comparison of the cloned sequences with the sequence of the *AT1G66690* gene showed that all cloned fragments correspond to the 3` fragment of the miR163-guided cleavage of the At1g66690 mRNA. Detailed analysis showed that all fragments start with nucleotides that correspond to the exact predicted cleavage site guided by miR163.

```
>gi|18408576|ref|NM_105340.1| Arabidopsis thaliana S-adenosyl-L-
methionine:carboxyl methyltransferase family protein (AT1G66690) mRNA,
complete cds
AAACTAAAGTTAAACAAGAAAAGAAGTTTGAGATAAAACAGAGAGATATATAATGACTACTACTCAAGATTGG
ATCATGATTGGTGGATATGGCCCTGAAAGTTACAACCAACAGTCATCGTATCAGAGAGCTTTGTTGGAAGCGG
CAAAGGACAAGATGACCGAGGCGATCTCCGCCAACCTCGACCTAGACTTGATTTCGAATCGCTTCATTGTAGC
GGATTTCGGTTGTGCAAGTGGACCTAACACTTTTGTGGCAGTCCAAAACATAATAGATGCCGTGGAAGAGAAG
TATCTTAGAGAAACCGGACAAAACCCGGAGGATAACATCGAGTTCCAAGTCCTCTTCAACGACTTAAGAATC
AATGACTTCAACACTCTCTTCCAGACACTCCCTCCGGGGAGAAGATACTTTAGCGCCGGGGTTCCTGGTTCCT
TCTTCAACCGTGTTCTTCCTAAGCAGAGTTTCCACATCGCAGTCATGAGTTACGCATTCCTTTTCACCTCCAA
AATCCCCAAAGGGATCATGGACCGCGACTCTCCGTTGTGGAACAAAGACATGCAGTGCACCGGGTTCAACCCC
GCTGTGAAGAAAGCCTATCTTGAGCAGTACTCTATTGACACCAAAAACCTTTTGGATGCTAGAGCTGAGGAGC
TCATGCCCGGGGGATTGATGTTGCTTTTAGGATCGTGTATGAGAGACGGAGTTAAGATGTCCGAGACCCTTAA
AGGAACAGTAATGGATTTTATTGGAGAATCTCTTAATGACCTTGCTCAAAAGGGTGTCACCGAACAGGAAAAG
GTGGACACTTTCAAAACCTCAATCTACTTTGCAGAACAAGGCGAGATAAGGCAAATCATTGAGGAGAATGGGA
```

AGTTCACAATCGAGGCTTTCGAGGATATCATTCACTCTAAGAATGAGTTTCCGCTAGACCCCAAGACATTGGC
CATCTCCTTCAAGGCCTTATATGGTGCTTTTATATCCGCACATTTTGGAATCGAAGTCATGAGGAAAGCCTTT
GAGCTTGTTGAGGTCAAGGCACGCGAACAGATTTCTCGCCTCCATAAGGTCAAACCCGGGATGCAATACCTCA
TCGTGCTTCGCAAGAACTGA

Figure 29 The coding sequence of the *At166690* gene. The hybridization site of the primer used for the first and second rounds of 5'RLM-RACE amplification is marked in cyan. The miRNA-complementarity site flanking the cleavage site is marked in red, and the first nucleotide found in the cloning products that corresponded to the At1g66690 transcript is underlined.

Additionally, we analyzed (by qPCR) a steady-state level of the *AT1G66690* gene transcript in wild-type plants as well as in the following miRNA biogenesis mutants: *hyl1-2, se-1, cbp20, cbp80, and cbc*. The obtained results from the qPCR are shown in Figure 25. We observed that the levels of mRNA in the miRNA-biogenesis mutants is higher compared to wild-type plants, which is in agreement with the decreased miR163 levels in these mutant plants (please compare Figures 26 and 30).

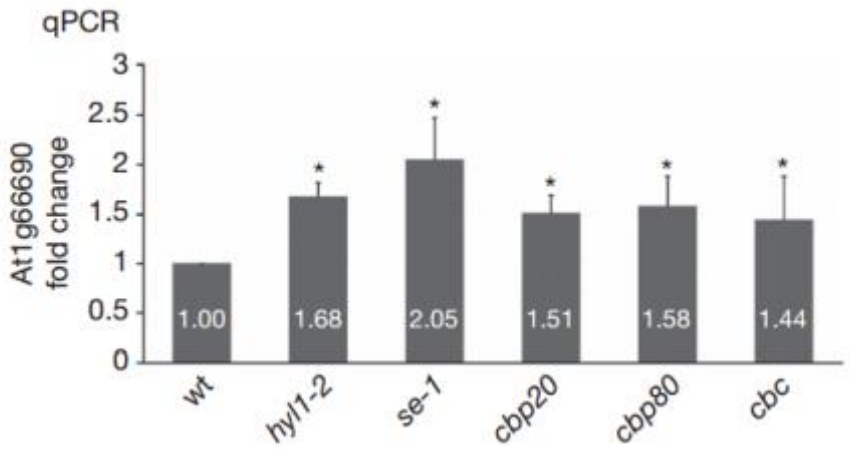


Figure 30 The levels of At1g666909 mRNA recorded in wild-type plants as well as in the following miRNA biogenesis mutants: *hyl1-2, se-1, cbo20, cbp80,* and *cbc*. Error bars indicate s.d. (n=3), and asterisks indicate a significant difference between the indicated sample and control wt plants (Mann-Whitney test, $p<0.05$).

7 Induction of miR163 depends on the functional splice sites in pri-miR163

MiR163 accumulation is induced by many agents, including salicylic acid, methyl jasmonate, and wounding [Ng *et al. 2011*]. The most effective inducer is a fungal elicitor alamethicin, which suggests that miR163 in *Arabidopsis thaliana* is involved in defense response after biotic stress [Ng *et al.* 2011]. We decided to compare the response of wild-type and IVSmut plants to infection caused by bacteria *Pseudomonas syringae* DC3000, which represents a model in molecular plant pathology [Hiranno & Upper 2000].

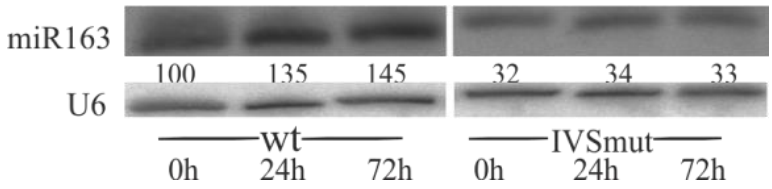


Figure 31 The results of Northern-blot analysis of mature miR163 in wild-type plants and the IVSmut transgenic line recorded after 0, 24, and 72 hours of infection with *P. syringae* DC3000. U6 snRNA served as a loading control.

Three-week-old plants grown in soil were sprayed with prepared bacteria (OD_{600} = 0.2 in 10mM $MgCl_2$ solution), and aerial parts of the sprayed plants were collected 24 or 72 hours after treatment. Total RNA was isolated from the collected samples and separated on a PAGE in the presence of 8M urea. The level of mature microRNA 163 recorded in the wild-type plants and the transgenic IVSmut line is shown is Figure 31. The analysis by Northern-blot hybridization showed that, after infection of wildtype plants, the level of miR163 increased by up to 135% after 24 hours of infection and by up to 145% after 72 hours of infection. In the IVSmut mutant, however, the level of miR163 did not change after 24 or 72 hours of infection.

We then decided to test the level of pri-miR163 in wild-type plants as well as in the IVSmut mutant plants after infecting them with *Pseudomonas syringae* DC3000. We therefore performed a RT-real-time PCR analysis using the same RNA samples as those used in the Northern-blot analysis. The obtained results are shown in Figure 32. The results showed that, in contrast to the level of mature microRNA163 in the wild-type plants, the primary transcripts of *MIR163* in the IVSmut plants decreased after infection. In contrast, the level of pri-miR163 increased in the IVSmut plants after bacterial infection. This observation suggests posttranscriptional regulation of miR163 biogenesis under biotic stress in wild-type *Arabidopsis thaliana* plants.

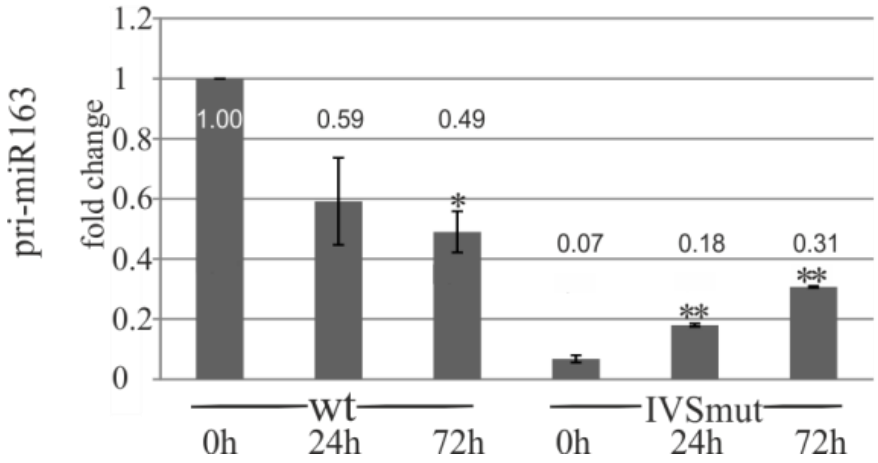


Figure 32 The level of the pri-miR163 recorded in wild-type and IVSmut plants after 0, 24, and 72 hours of infection with *P. syringae* DC3000. Error bars indicate s.d.(n=3). One asterisk indicates a significant difference between the indicated sample and wild-type plants at the 0h time point, and the double asterisk indicates significant difference between the indicated sample and the IVSmut plants at the 0h time point (Mann–Whitney test, $p<0.05$). wt, wild type.

Moreover, we decided to analyze (by qPCR) the level of At1g66690 transcripts in wild-type plants and in IVSmut mutant plants after infecting them with *Pseudomonas syringae* DC3000. The obtained results of the qPCR analysis are shown in Figure 33. The qPCR results showed that the level of the At1g66690 transcript decreased after bacterial infection of the wild-type plants. In contrast, the level of the At1g66690 transcript increased after bacterial infection of the IVSmut mutant plants. These observations showed that the regulation of miR163 biogenesis under biotic stress relies on functional splice sites in the *MIR163* gene.

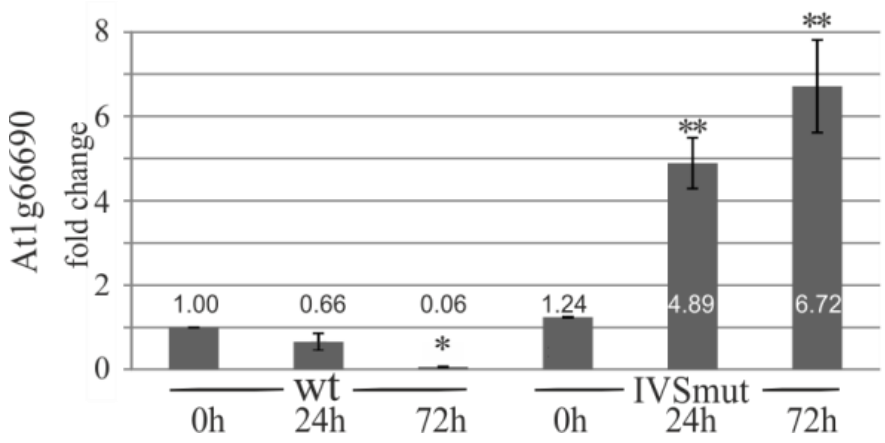


Figure 33 The level of mRNA in At1g66690 recorded in wild-type and IVSmut plants after 0, 24, and 72 hours of infection with *P. syringae* DC3000. Error bars indicate s.d.(n=3). One asterisk indicates a significant difference between the indicated sample and wild-type plants at the 0h time point, and the double asterisk indicates significant difference between the indicated sample and the IVSmut plants at the 0h time point (Mann–Whitney test, $p<0.05$). wt, wild type.

Transcriptional regulation of biogenesis of microRNAs

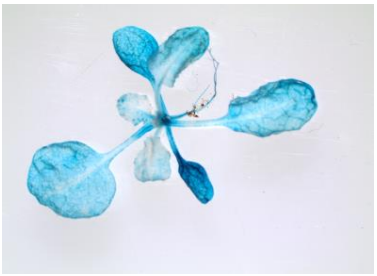
It was shown that pri-miRNAs accumulate in *hyl1-2* mutants as compared to wild-type plants [Vazquez *et al.* 2004a, Szarzyńska *et al.* 2009]. Additionally, data obtained by Szarzyńska and her colleagues showed that an unspliced isoform of several pri-miRNAs accumulates in the *hyl1-2* mutant. Moreover, it is well known that intron removal by spliceosme is a co-transcriptional event, which means that splicing occurs before nascent RNA synthesis is completed [Beyer & Osheim 1988, Neugebauer 2002]. Additionally, it was shown that the HYL1 protein needs to be dephosphorylated by a CPL1 protein that is also able to dephosphorylate a CTD of the RNA POL II at serine 5 residue [Manavella *et al.* 2012, Koiwa *et al.* 2004]. These observations prompted us to investigate if a HYL1 protein is involved in the transcription of *MIR* genes.

1 Analysis of GUS reporter lines of *Arabidopsis thaliana* in a *hyl1-2* background.

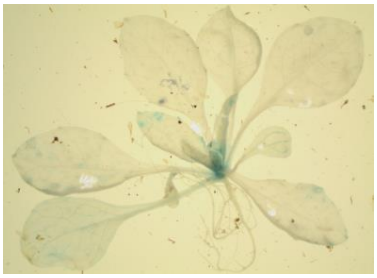
To test a hypothesis that the HYL1 protein in involved in transcription of *MIR* genes, we decided to examine the effect of the absence of the HYL1 protein on the expression of a GUS reporter gene under control of an *MIR* gene promoter. Previously described transgenic lines pMIR393A::GUS and pMIR393B::GUS were crossed into a *hyl1-2* background [Parry *et al.* 2009]. Seeds from self-pollinated F1 progeny after crossing were sown in soil and grown in a long-day photoperiod in a Sanyo chamber until the siliques matured. Selection of desired plants among F2 plants was based on phenotypic traits (*hyl1-2* mutation) and on a PCR-based method (the presence of a GUS reporter gene). Next, the intensity of GUS staining was examined in the *hyl1-2* mutant background and was compared to that of the Col-0 background. Results of the intensity of GUS staining are shown in Figure 34.

A)

pMIR393A::GUS



Col0 background



hyl1-2 background

B)

pMIR393B::GUS

Col0 background

hyl1-2 background

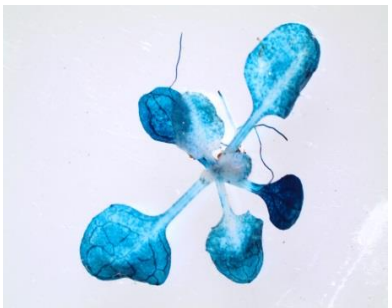
Figure 34 Expression of pMIR393A::GUS (A) or pMIR393B::GUS (B) in seedlings in wild-type or *hyl1-2* backgrounds.

A lack of intensity of the GUS staining signal was observed in the *hyl1-2* mutant background, compared to that of the Col-0 background. This data suggests another HYL1 role. Besides its role in pri-miRNA processing HYL1 may also regulate (or be involved in) the transcription of *MIR* genes, at least in the cases of *MIR393A* or *MIR393B*.

We were also interested if the effect of *hy1l-2* mutation on the expression of a GUS reporter gene is restricted to *MIR* genes. We decided to use transgenic lines where a GUS reporter gene is under control of a protein-coding gene promoter. We crossed previously described transgenic lines pTIR1::GUS, pAFB1::GUS, and pAFB2::GUS into a *hyl1-2* background [Parry *et al.* 2009]. Desired F2 plants were selected in the same way as described above. Next, the intensity of GUS staining was examined in the *hyl1-2* mutant background and was compared to that of the Col-0 background. Results in intensity of GUS staining are shown in Figure 35.

A)

pTIR1::GUS



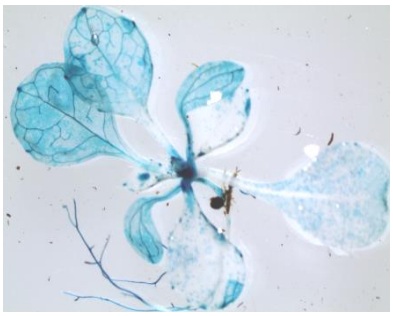
Col0 background

hyl1-2 background

B)

pAFB1::GUS

Col0 background



hyl1-2 background

No obvious changes in the intensity of the GUS staining signal were observed in the *hyl1-2* mutant background when compared to that of the Col-0 background. This data suggests that, besides its role in pri-miRNA processing, HYL1 may also regulate or be involved in the transcription of *MIR* genes, at least in the cases of the *MIR393A* or *MIR393B*.

2 Analysis of RNA Pol II distribution in a *hyl1-2* mutant at *MIR* genes.

If the HYL1 protein is involved in the transcription of *MIR* genes, the occupancy of RNA Pol II in the promoter region or the *MIR* gene's body might be changed in the *hyl1-2* mutant (as compared to wild-type plants). To test the occupancy of the RNA Pol II on *MIR* genes, a chromatin immunoprecipitation (ChIP) assay using antibody against total RNA Pol II was conducted. DNA samples obtained after immunoprecipitation from the *hyl1-2* and wild-type plants were analyzed using qPCR. ChIP analysis was done for the *MIR393A* and *MIR393B* genes. Results of the ChIP analysis are shown in Figure 36 for *MIR393A* and Figure 37 for *MIR393B*.

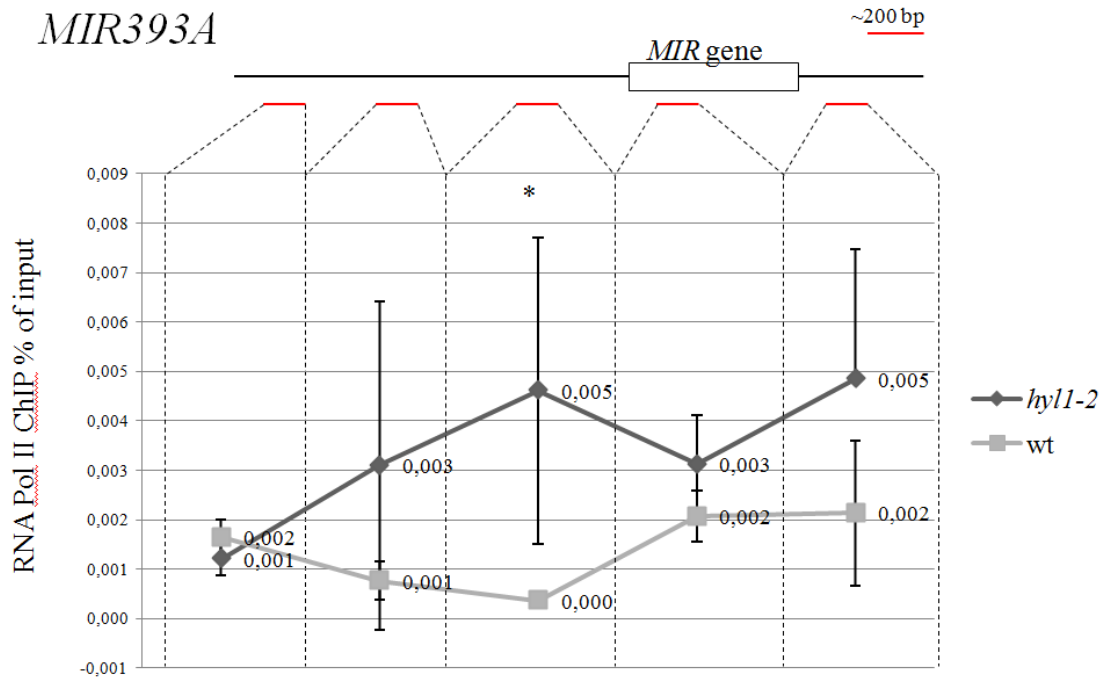


Figure 36 The occupancy of total RNA Pol II at the *MIR393A* locus using ChIP. The region marked with an asterisk represents the statistically significant enrichment of RNA Pol II in *hyl1-2* as compared to wild-type plants. Schematic gene structure is shown, and red lines show the amplified regions. Error bars represent the SD of three independent biological replicates.

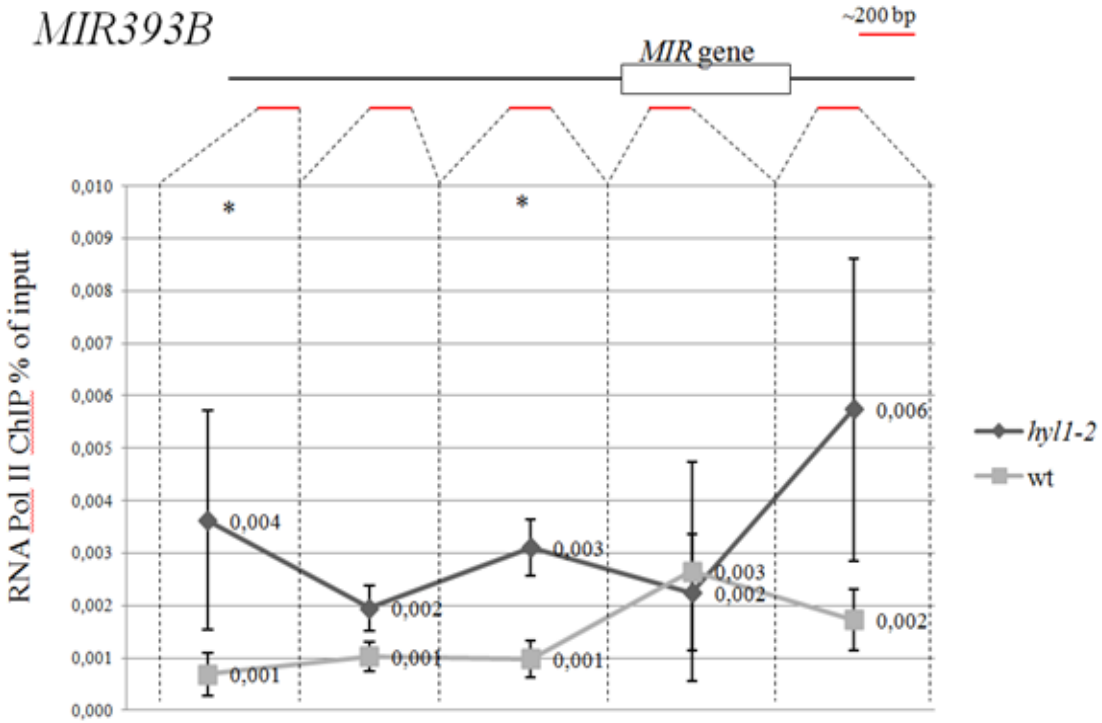


Figure 37 The occupancy of total RNA Pol II at the *MIR393B* locus using ChIP. Regions marked with an asterisk represent the statistically significant enrichment of RNA Pol II in the *hyl1-2* as compared to wild-type plants. Schematic gene structure is shown, and red lines show the amplified regions. Error bars represent the SD of three independent biological replicates.

ChIP experiments showed an accumulation of total RNA Pol II at the region approximately 200 base pair upstream from the transcription initiation site (TSS) for *MIR393A* and *MIR393B* in the *hyl1-2* mutant. The higher occupancy of RNA polymerase II in these regions may suggest that, in the *hyl1-2* mutant: 1) transcription of *MIR393A/MIR393B* genes is higher; or 2) transcription transition from initiation to elongation is deregulated. The results presented in the previous chapter showed that transcription of the GUS reporter was lower in a *hyl1-2* mutant when compared to wild-type plants (Figure 29). Altogether, these experiments suggest that the accumulation of RNA Pol II in *hyl1-2* is responsible for the lower transcription of *MIR393A/MIR393B* genes.

Zhang and his colleagues showed that MYB-transcription-factor CDC5 interacts with promoters of *MIR* genes and RNA Pol II [Zhang *et al.* 2013]. In our experiments, we decided to also include a set of *MIR* genes that were tested by Zhang and compare the occupancy of RNA pol II (~200 bp upstream of TSS) of different *MIR* genes in *hyl1-2* and wild-type plants. The results of the ChIP analysis are shown in Figure 38.

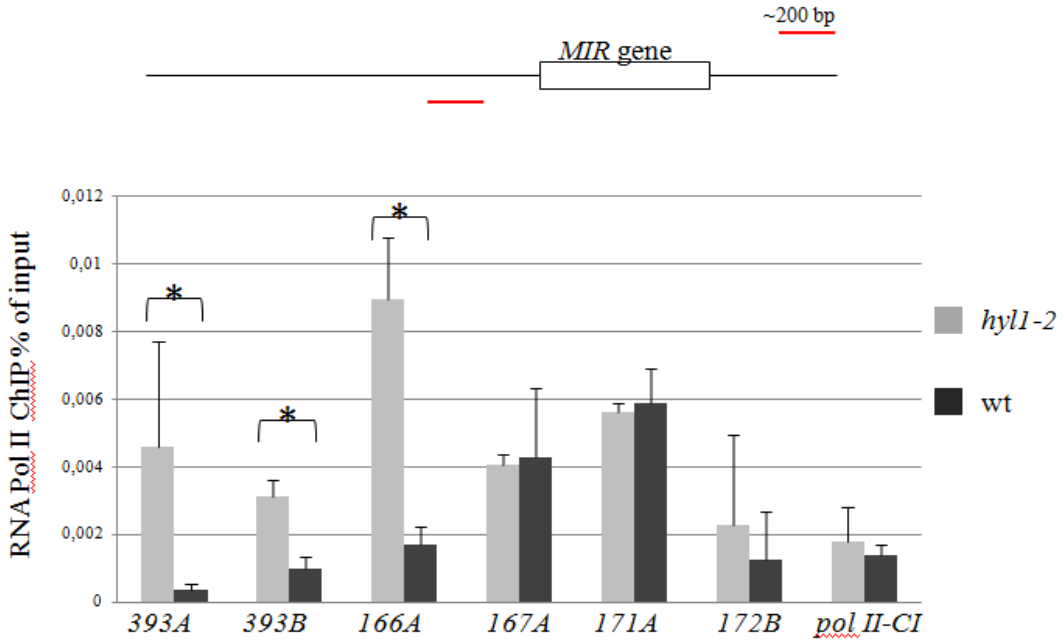


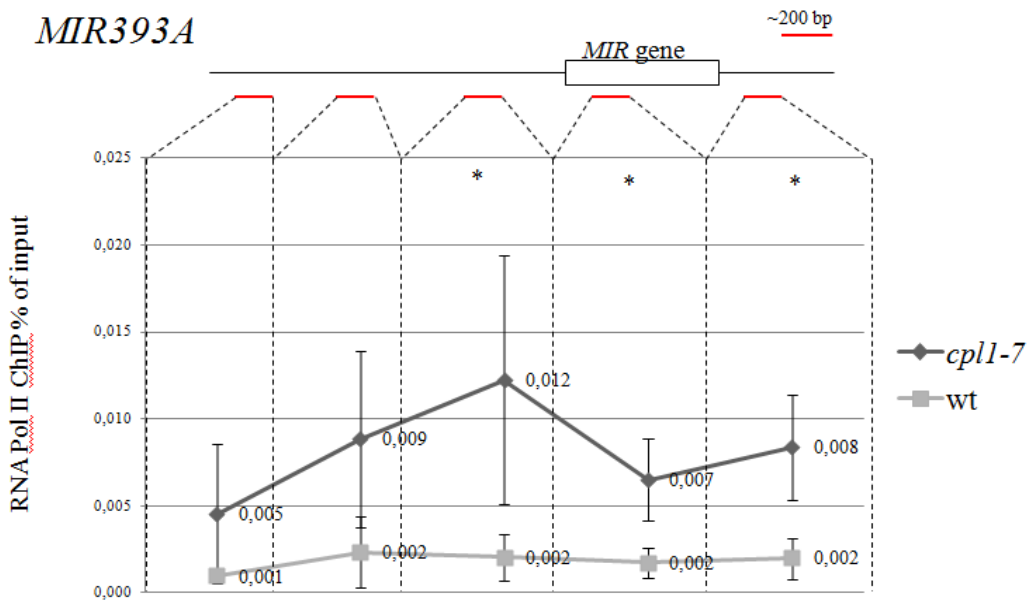
Figure 38 The occupancy of total RNA Pol II at promoter regions (~200 bp upstream of TSS) of different *MIR* genes. Asterisk indicates a statistically significant difference between the *hyl1-2* and wild type samples. Schematic gene structure is shown and red lines shows amplified region. Error bars represent SD of three independent biological replicates. As an additional control, primers amplifying an intergenic region (pol II-CI) were used.

Our ChIP experiments showed the enrichment of RNA Pol II at the promoter region of *MIR393A*, *MIR393B,* and *MIR166A* in the *hyl1-2* mutants as compared to wild-type plants. Moreover, results presented have shown that a GUS reporter that was under control of an *MIR* gene promoter is expressed at a lower level when HYL1 is missing, so the higher occupancy of RNA polymerase II suggests that transcription transition from initiation to elongation is deregulated. However, for *MIR167A*, *MIR171A*, and *MIR172B*, the ChIP experiments did not show any statistically significant changes of RNA Pol II occupancy at the promoter regions of the *hyl1-2* mutant. It is possible that deeper insight into the *MIR* genes promoter's sequences will shed a light on HYL-dependence RNA Pol II accumulation in the promoter regions.

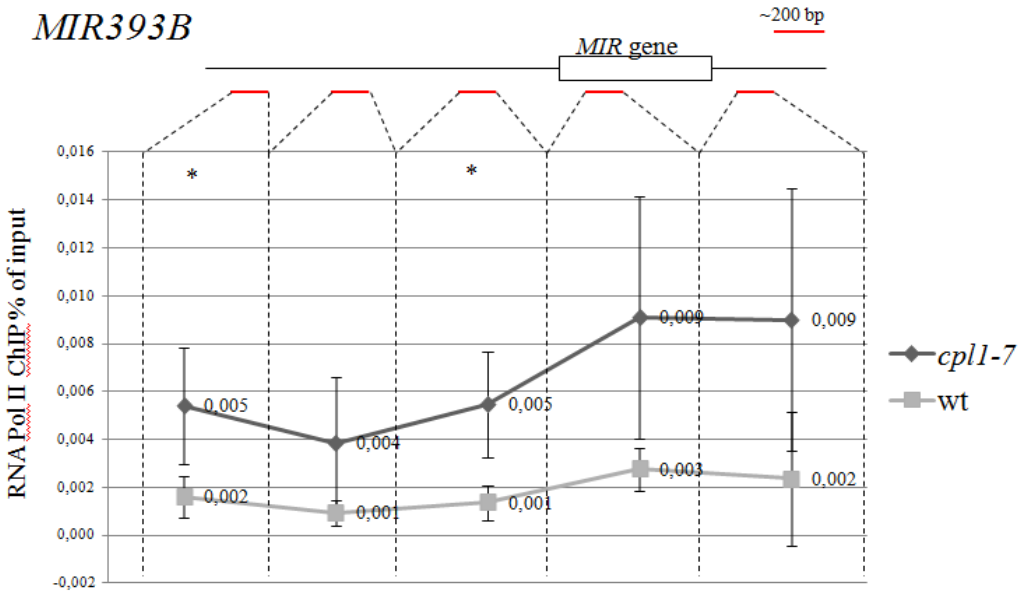
3 Analysis of RNA Pol II distribution in the *cpl1-7* mutant at *MIR* genes.

Previously, it was shown that CPL1 is able to dephosphorylate a CTD of RNA POL II at serine 5 residue [Koiwa *et al.* 2004]. Additionally, it was also shown that the HYL1 protein needs to be dephosphorylated for its optimal activity, and hypophosphorylation of HYL1 is maintained by the CPL1 protein. Morever, CPL1 physically interacts with the HYL1 protein [Manawella *et al.* 2012, Jeong *et al.* 2013]. These reports prompted us to analyze the occupancy of total RNA Pol II by ChIP in *cpl1-7* and wild-type plants. For the qPCR in the ChIP analysis, we used the same set of primers as were used in ChIP analysis in the previous chapter. The results for all of the tested regions are presented in Figure 39.

A)



B)



C)

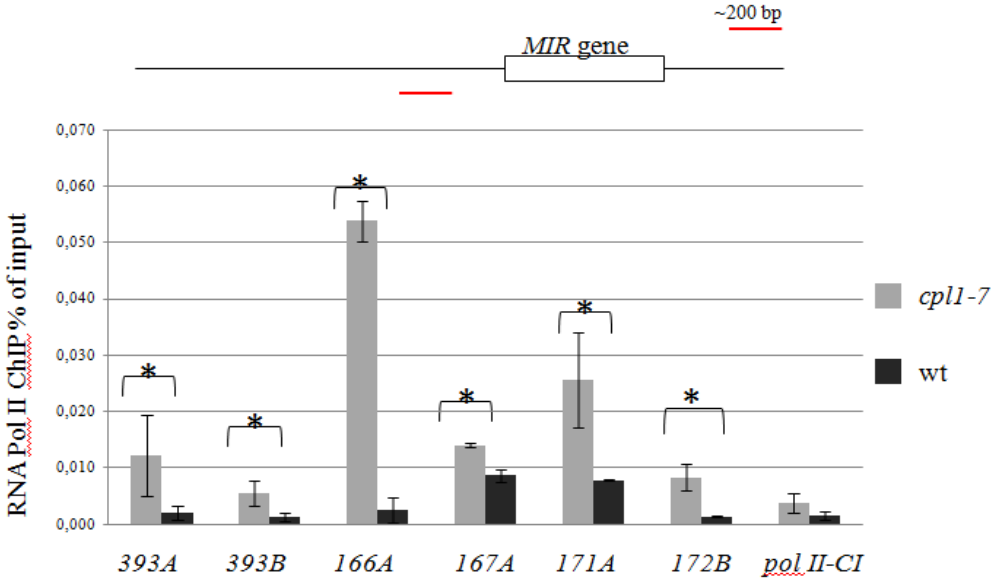


Figure 39 The occupancy of total RNA Pol II at the MIR*393A* locus *(A)*, the *MIR393B* locus (B), or at promoter regions (~200 bp upstream of TSS) of other studied *MIR* genes (C). An asterisk indicates a statistically significant difference between the *cpl1-7* and wild-type plants. Gene structure is shown, and red lines show the amplified region. Error bars represent the SD of three independent biological replicates. As an additional control in C, primers amplifying an intergenic region (pol II-CI) were used.

In the case of the *MIR393A* locus, the ChIP experiments showed an accumulation of total RNA Pol II at the region approximately 200 base pair upstream from the transcription initiation site, at the gene body, and also downstream of the *MIR393A* gene.

In the case of the *MIR393B* locus, accumulation of total RNA Pol II 600 bp and 200 bp upstream of the TSS was observed. Moreover, analysis of the promoter regions of the *MIR166A*, *MIR167A*, *MIR171A*, and *MIR172B* genes showed accumulation of RNA Pol II in the *cpl1-7* mutant as compared to samples from the wild-type plants. CPL1 affects RNA Pol II accumulation in the case of analyzed *MIR* genes promoters while HYL1 is not.

4 Analysis of pri-miRNA163 transcription and processing under the control of different promoters in *hyl1-2* or wild-type plants.

To evaluate the importance of the HYL protein in the initiation of the transcription of RNA Pol II at *MIR* genes, we constructed and introduced the *MIR163* gene under different promoters in the *mir163-2* mutant or *mir163-2*x*hyl1-2* double mutant background. To obtain a double mutant, a cross between *mir163-2* and *hyl1-2* mutants was performed. After the crossing, the F2 progeny was used for the selection of the double *mir163-2*x*hyl1-2* mutant. Four constructs where generated:

1) pMIR163::MIR163 is a construct representing a native-gene version of *MIR163*.

2) p35S::MIR163 is a construct where the native promoter of *MIR*163 was replaced with the 35S CaMV promoter.

3) pACT2::MIR163 is a construct where the native promoter of *MIR163* was replaced with the promoter of the *ACT2* gene. The promoter region represents ~2000 bp upstream of the TSS of the At3G8780 (ACT2) gene.

4) pGAPDH::MIR163 is a construct where the native promoter of *MIR163* was replaced with the promoter of the *GAPDH* gene. The promoter region represents ~2000 bp upstream of the TSS of the At1G13440 (GAPDH) gene.

Schematic representation of these constructs can be found in Figure 40.

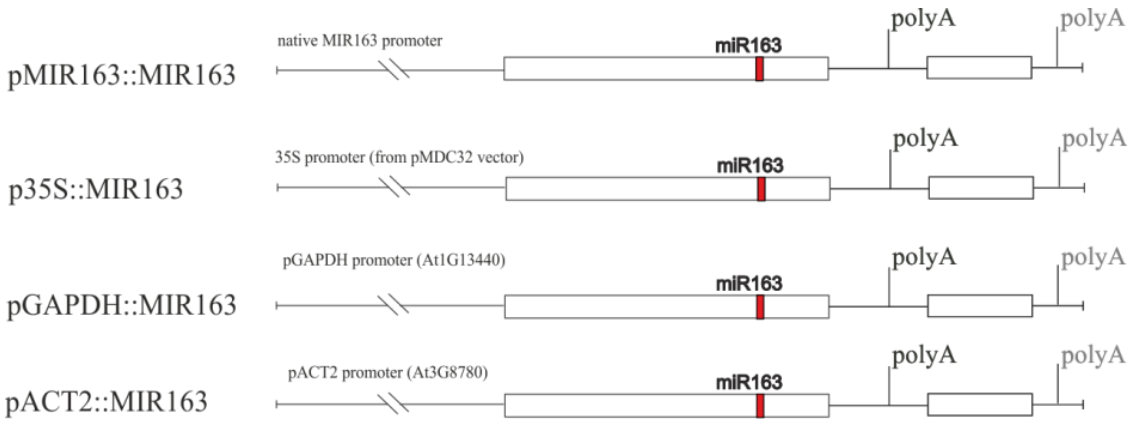


Figure 40 Schematic representation of the *MIR163* gene variants. MicroRNA163, position of proximal, and distal poly(A) sites are marked.

A construct representing a native gene of the *MIR163* (pMIR163::MIR163) was amplified by PCR using A01 and A02 primers and genomic DNA from wild-type plants as a template (for primers, see table from Chapter 3 in the Materials section). Next, the amplicon was excised from agarose gel and cloned to the pENTR D/TOPO plasmid, followed by an LR recombination reaction with the pMDC99 binary vector. For p35S::MIR163 construct preparation, a PCR was conducted, using specific primers and genomic DNA from wild-type plants as a template. The primers were design to the amplified *MIR163* gene body (for primers, see table from Chapter 3 in Materials). Next, the amplicon was excised from the agarose gel and cloned to the pENTR D/TOPO plasmid using NotI and AscI restriction sites. The pENTR D/TOPO with cloned *MIR163* gene body was used in an LR recombination reaction with binary vector pMDC32 (a strong 35S promoter from the CaMV virus is included in this vector). For the pACT2::MIR163 or the pGAPDH::MIR163 constructs, the promoter regions of *ACT2* or *GAPDH* were first cloned *via* HindIII and AscI restriction sites to the pMDC99 binary vector. Primers A21 and A23 were used for amplification of the pACT promoter, and primers A23 and A24 were used for the pGAPDH promoter. Genomic DNA from wild-type plants was used as a template. Next, the pENTR D/TOPO with cloned *MIR163* gene body was used in an LR recombination reaction with the previously prepared pMDC vectors (pMDC-pACT2 or pMDC-pGAPDH).

The obtained binary vectors were introduced by electroporation to the *Agrobacterium tumefaciens* AGL-1 strain. The prepared strains of AGL-1 bacteria were used for transformation of the *mir163-2* or *mir163xhyl1-2* mutants by the floral-dip method [Clough & Bent 1998].

After transformation, the plants were grown in a long-day photoperiod condition in a Sanyo chamber until the siliques turned yellow and then brown. Collected seeds (T1 generation of transformed plants) were sieved on a ½ MS medium supplemented with hygromycin. The seeds were incubated in Petri dishes in a Sanyo chamber for two weeks. Next, the green seedlings were transferred to soil and grown in a short-day photoperiod. Arbitrarily selected samples (at least five independent transgenic lines) from each T1 generation were grown for five weeks.

RNA was isolated separately from each of the selected plants using a TRIzol reagent. Next, after DNase I treatment, an RT-qPCR was conducted. Results of the RT-qPCR are shown in Figure 36. Analysis of the transgenic plants showed that, when a native promoter of *MIR163* was used, pri-miR163 accumulated in the *mir163-2xhyl1-2*

background as compared to the *mir163*-2 mutant background. It was observed that the majority of pri-miRNA accumulated in the *hyl1-2* mutant, and pri-miR163 was one of them [Vazquez *et al*. 2004a, Szarzyńska *et al*. 2009]. Analysis of transgenic plants where pri-miR163 is under control of the 35S promoter showed that, although the expression level of pri-miR163 is very high in transgenic plants, it is not significantly different between the *mir163-2* and *mir163-2xhyl1-2* mutants' backgrounds. The same results can be observed when pri-miR163 was under control of pACT or pGAPDH promoters.

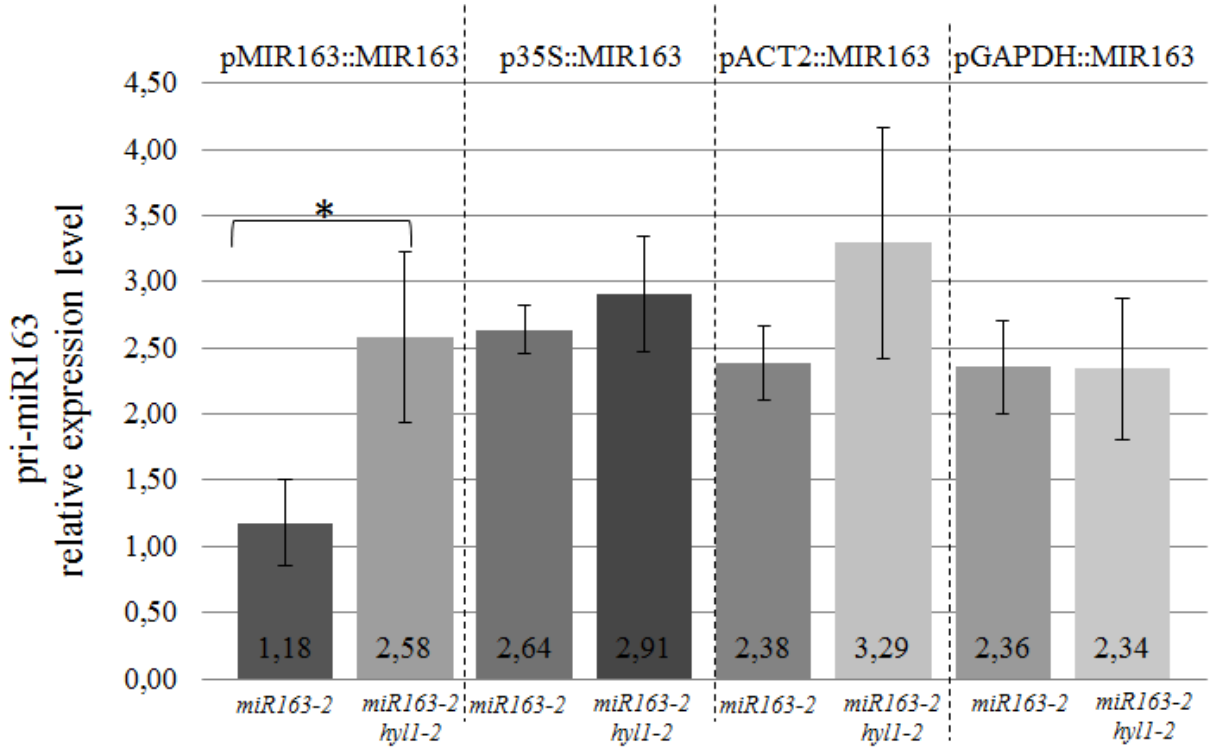


Figure 41 The results of the qPCR analysis of the T1 generation of transgenic lines carrying *MIR163* gene constructs under control of different promoters. Asterisks indicate statistically significant differences between the indicated samples. Error bars represent the SD of five independent transgenic lines.

To analyze the level of mature miR163 in the transgenic line, Northern-blot hybridization was performed. The results of the analysis are shown in Figure 42.

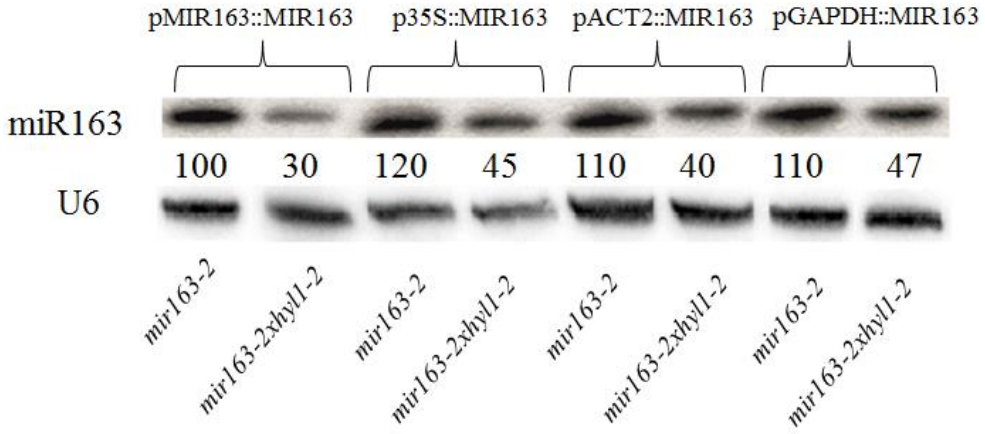


Figure 42 Results of Northern-blot analysis of the *mir163-2* and *mir163-2xhyl1-2* transgenic lines carrying the pMIR163::MIR163, p35S::MIR163, pACT2::MIR163 and pGAPDH::MIR163 constructs. U6 snRNA served as a loading control.

The obtained results showed that the HYL1 protein is required for the accumulation of the proper level of miR163. In all *MIR163* gene variants tested, the level of mature miR163 was lower in the *hyl1-2* background. However, using Northern-blot hybridization, we cannot discriminate between two processes: the transcription and maturation of pri-miRNAs.

Although, we know that RNA POL II accumulates in the *hyl1-2* background at a region upstream from the TSS, and this leads to decreased transcription from *MIR393A* and *MIR393B loci* (please check Figures 38C and 34). Moreover, only in the case of the *MIR163* gene variant with native promoter, pri-miRNA levels are different between the wild-type and *hyl1-2* backgrounds. These observations suggest that HYL1 is involved in the transcription of MIR genes.

5 Analysis of pri-miRNA161 transcription and processing under control of different promoters in a *hyl1-2* or wild-type background.

To evaluate whether involvement of the HYL1 protein in transcription is not specific to the *MIR163* gene, we tested another *MIR* gene. For this purpose, we have chosen the previously described *MIR161* gene. We constructed the *MIR161* gene under control of a different promoter. We used the same promoters as in the case of *MIR163*: native, 35S from CaMV, pACT2, and pGAPDH. Additional modification in the constructs was replacing the miR161 and miR161 star sequences with miR163 and mirR63 stars, respectively. There is no *mir161* null mutant, and there are no available *Nicotiana* plants

without a HYL1 protein. This replacement in the *MIR161* gene constructs will allow us to use the *mir163-2* mutant or *mir163-2*x*hyl1-2* double-mutant plants for transformation. After the introduction of the prepared constructs into the previously mentioned mutant background, we will be able to follow microRNA production from the *MIR161/163* gene constructs (this should not interfere with endogenous *MIR161* gene activity). Four constructs where generated (in all constructs, the miR161 and miR161 star sequences were replaced with miR163 and miR163 stars, respectively):

1) pMIR161::MIR161/163 is a construct representing a native-gene version of *MIR161*.

2) p35S::MIR161/163 is a construct where the native promoter of *MIR161* was replaced with the 35S CaMV promoter.

3) pACT2::MIR161/163 is a construct where the native promoter of *MIR161* was replaced with the promoter of the *ACT2* gene. The promoter region represents ~2000 bp upstream of the TSS of the At3G8780 (ACT2).

4) pGAPDH::MIR161/163 is a construct where the native promoter of *MIR161* was replaced with the promoter of the *GAPDH* gene. The promoter region represents ~2000 bp upstream of the TSS of the At1G13440 (GAPDH).

Constructs can be found in Figure 43.

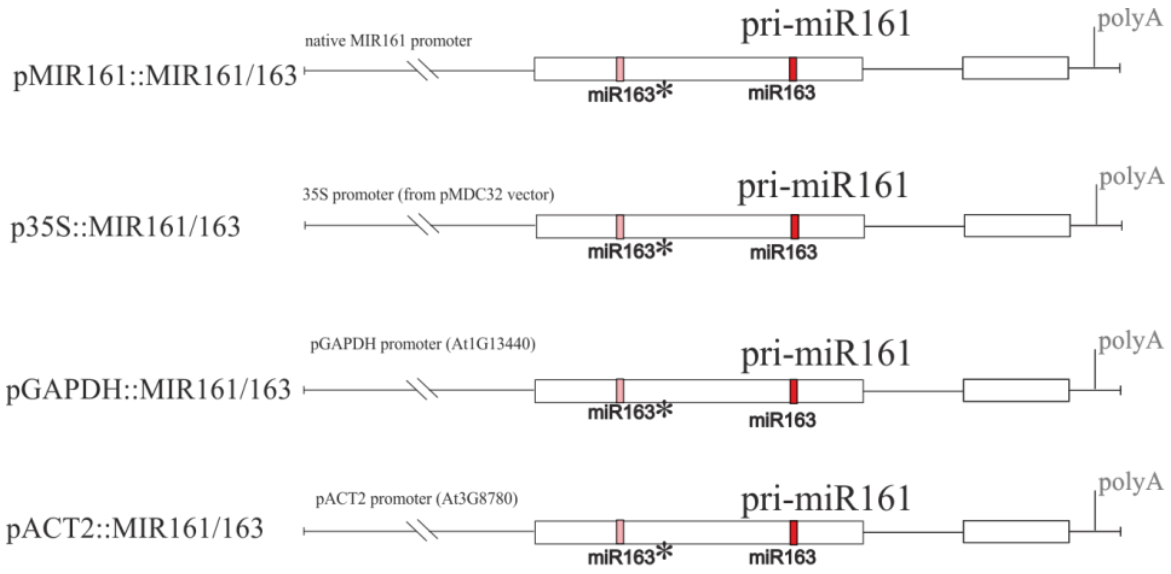


Figure 43 Schematic representation of the *MIR161/163* gene variants. MicroRNA and microRNA star positions are marked.

The construct representing the native gene of *MIR161* (pMIR161::MIR161/163) was amplified by a PCR using A11 and A12 primers and genomic DNA from wild-type plants as a template (for primers, see table from Chapter 3 in Materials). Next, the amplicon was excised from the agarose gel and was cloned to the pCR8 plasmid. The prepared pCR8 plasmid with the *MIR161* gene was used as a template in two PCRs mutagenesis reaction (replacement of microRNA and microRNAstar sequences), which resulted in the final plasmid pCR8_pMIR161::MIR161/163. The plasmid was used in an LR recombination reaction with a pMDC99 binary vector. For preparation of three other constructs (see Figure 43), a PCR was conducted using plasmid pCR8_pMIR161::MIR161/163 as a template and primers amplifying the *MIR161* gene body. Next, the amplicon was excised from the agarose gel and cloned to the pCR8 plasmid, which resulted in the final plasmid pCR8 MIR161/163. The pCR8_ MIR161/163 plasmid was used in an LR recombination reaction with binary vector pMDC32, pMDC-pACT2, or pMDC-pGAPDH.

The obtained binary vectors were introduced by electroporation to the *Agrobacterium tumefaciens* AGL-1 strain. The prepared strains of AGL-1 bacteria were used for transformation of *mir163-2* or *mir163xhyl1-2* mutants by the floral-dip method [Clough & Bent 1998].

After transformation, the plants were grown in a long-day photoperiod condition in a Sanyo chamber until the siliques turned yellow and then brown. Collected seeds (T1 generation of transformed plants) were sieved on a ½ MS medium supplemented with hygromycin. The seeds were incubated in Petri dishes in a Sanyo chamber for two weeks. Next, the green seedlings were transferred to soil and grown in a short-day photoperiod. Arbitrarily selected samples (at least five independent transgenic lines) from each T1 generation were grown for five weeks.

RNA was isolated separately from each of the selected plants using a TRIzol reagent. Next, after DNase I treatment, RT-qPCR was conducted. Results of the RT-qPCR are shown in Figure 44. Analysis of the transgenic plants showed that, when the native promoter of *MIR161* was used in a construct, the expression level of pri-miR161/163 was the same in the *mir163-2* and *mir163-2xhyl1-2* mutant backgrounds. However, Szarzyńska and her colleagues described that pri-miR161 accumulates in the *hyl1-2* mutant [Szarzyńska *et al*. 2009]. This discrepancy may be connected with the fact that MIR161/163 gene constructs are artificial for *Arabidopsis thaliana*. Analysis of transgenic plants where *MIR161/163* is under control of 35S, pACT, or pGAPDH promoters also

showed the same expression level of pri-miR161/163 in *mir163-2* and *mir163-2xhyl1-2* mutant backgrounds.

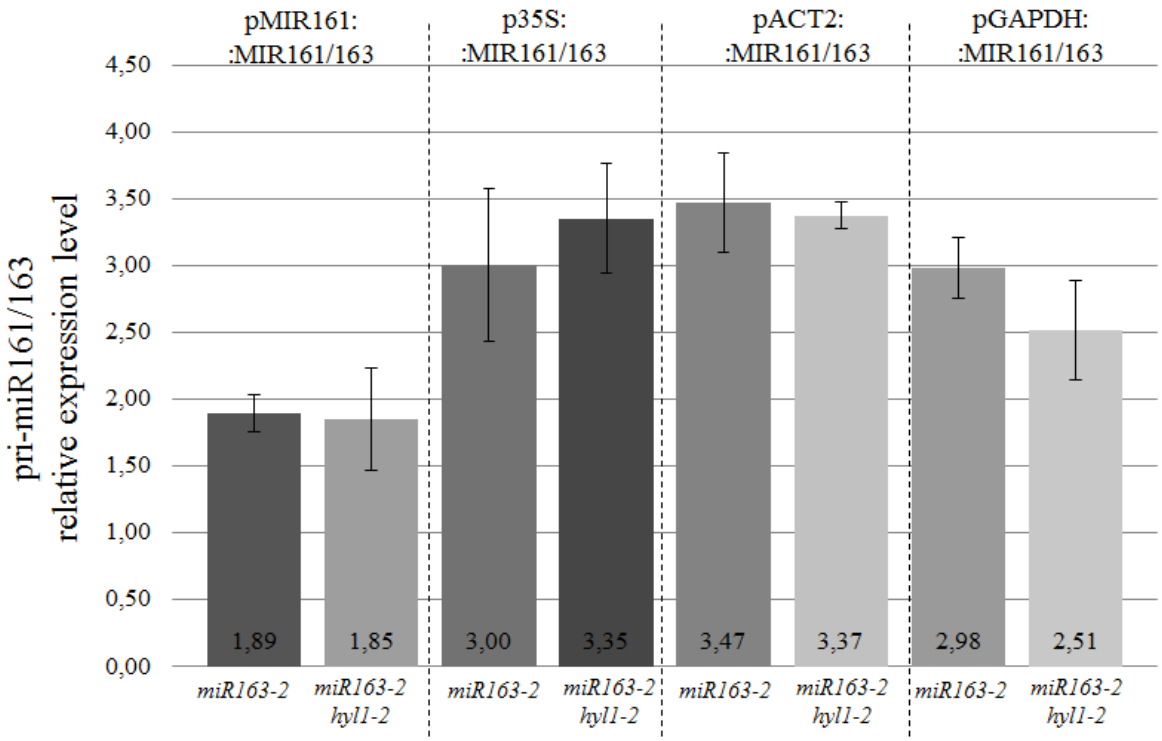


Figure 44 The results of qPCR analysis of the T1 generation of transgenic lines carrying *MIR161/163* gene constructs under control of different promoters. Error bars represent the SD of five independent transgenic lines.

To analyze the level of mature microRNA rising from the *MIR161/163* gene in the transgenic line, Northern-blot hybridization with a probe against microRNA163 was performed. The results of the analysis are shown is Figure 45

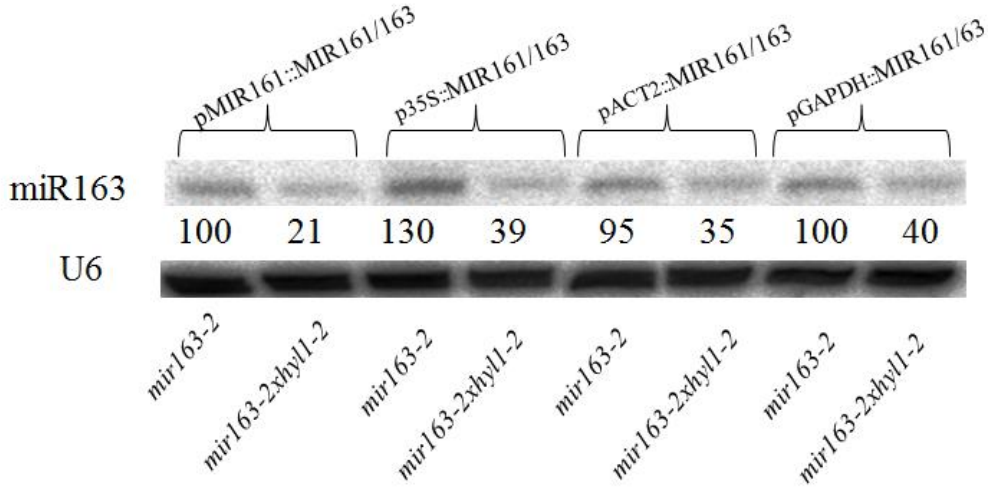


Figure 45 Results of Northern-blot analysis of the *mir163-2* and *mir163-2xhyl1-2* transgenic lines carrying the pMIR161::MIR161/163, p35S::MIR161/163, pACT2::MIR161/163 and pGAPDH::MIR161/163 constructs. U6 snRNA served as a loading control.

The obtained results showed that the HYL1 protein is required for accumulation of the proper level of mature microRNA generated from pri-miR161/163. In all of the *MIR161/163* gene variants tested, in the *hyl1-2* background level of mature microRNA was decreased. However, using Northern-blot hybridization, we cannot discriminate between two processes: the transcription and maturation of pri-miRNAs.

DISCUSSION

1 Post-transcriptional regulation of the biogenesis of microRNAs

Generally, *MIR* genes can be classified as independent transcriptional units (intergenic) or as microRNA residing within the introns of host-genes (intronic). MiRNAs encoded within the intergenic *loci* are transcribed from their own promoters. Intronic miRNAs are located within the introns of host-genes and are thought to be derived from the introns of their host genes. The majority of plant microRNAs are encoded by *MIR* genes representing independent transcriptional units [Rodriguez *et al.* 2004; Xie *et al.* 2005b, Szarzyńska *et al*. 2009, Szweykowska-Kulińska *et al*. 2013]. Typical plant *MIR* genes have an exon-intron structure that resembles protein-coding genes. Additionally, it was reported that introns localize 3` to the microRNA stem-loop structure in the most cases of intergenic *MIR* genes [Szarzyńska *et al.* 2009]. It has been thought that the region outside the microRNA stem-loop is dispensable and has no regulatory role. However, introns within pri-miRNA undergo complex alternative splicing, providing a possible mechanism for the regulation of miRNA processing [Yan *et al.* 2012, Sobkowiak *et al.* 2012]. A possible function or connection between the splicing and processing of *MIR* genes that are located in intergenic regions has not been investigated in detail. We decided to test a connection between the dicing and splicing of plant pri-miRNAs. In this work, we have presented experiments that revealed that both introns and at least U1 snRNP that is a component of the spliceosome are important for the regulation of microRNA levels.

Interplay between splicing machinery and microprocessor affects microRNA biogenesis

To evaluate the significance of introns in pri-miRNAs and their role in the biogenesis of mature microRNAs, we constructed and introduced original or mutated *MIR163* gene variants under a native *MIR163* promoter in the *mir163-2* mutant background. Analysis of transformants revealed that an intron-containing transgene fully restored the level of microRNA163. By contrast, intron-less transgene expressed approximately a three-times-lower level of microRNA than in wild-type plants. Similar transient expression experiments performed in tobacco leaves and with *MIR161* gene variants have shown a strong downregulation of microRNA expression from intron-less variants. Moreover, Schwab and her colleagues also expressed intron-less variants of *MIR163* and *MIR172A* (another intron-containing *MIR* gene) in *Arabidopsis thaliana* and *Nicotiana benthamiana* plants [Schwab *et al.* 2013]. Results obtained by Schwab and her

colleagues clearly demonstrated that higher levels of mature microRNA are generated from intron-containing pri-miRNAs. Additionally, they used three introns from protein-coding genes (UBQ, TRP1, and TCH3) and introduced them at the 3' end of the intron-less *MIR163* variant. Results showed that UBQ or TCH3 introns increased microRNA 163 levels [Schwab *et al*. 2013]. Our observations (as well as those of Schwab *et al.)* imply that introns within pri-miRNA are important for boosting the accumulation of microRNAs. Additionally, the intron nucleotide sequence is not important for a stimulatory effect. However, it seems that the localization of introns with respect to the miRNA stem-loop structure is important. Schwab and her colleagues showed that the presence of a 5' intron had only a minor effect on mature microRNA level in the case of the *MIR17B* [Schwab *et al.* 2013].

The formation of an active spliceosome is initiated by recognition and binding by U1 snRNP to the donor splice site (5' splice site). Mutation of the 5` splice site in *MIR161* and *MIR163* had a strong effect on decreasing the accumulation of a given microRNA. However, in the study reported by Schwab and her colleagues, mutation of the 5` spice site did not affect microRNA accumulation [Schwab *et al.* 2013]. These differences might be due to the nature of mutation introduced to inactivate the 5` splice site. In our work, we mutated six nucleotides within the 5' splice site, while Schwab *et al.* introduced only mutations of the two first nucleotides of the given intron. It cannot be excluded that the mutation of these two particular nucleotides did not completely abolish U1 snRNP binding to the 5` splice site. Moreover, the engagement of splicing factors to the transcribed RNA might be affected by the type of promoter used in constructs. In the case of experiments carried out in this PhD dissertation, variants of the *MIR163* gene were under a native promoter, while Schwab *et al.* used a strong 35S promoter from CaMV in their research. However, it is clear from both of our studies that the plant miRNA microprocessor interplays with the splicing machinery. Experiments provide a suggestion that certain proteins that form U1 snRNP can interact with one or more of the main players of microRNA biogenesis in plants.

In the selected *Arabidopsis thaliana* null mutants of SR protein genes that comprise a conserved family of pre-mRNA splicing factors, the level of some (miR161, miR163, and miR171) intron-containing microRNAs was decreased as compared to wild-type plants. As previously reported, at least two SR proteins (SRZ-21 and SRZ-22) interact with the 70K protein that is a component of U1 snRNP [Golovkin & Reddy 1998]. Additionally, our unpublished data indicates that the SERRATE protein (one of the main players of

microRNA biogenesis in plants) can interact with certain U1 snRNP proteins. To our knowledge, the findings presented in this dissertation and the study published by Schwab and her colleagues are the first reports showing that splicing (in the case of intergenic *MIR* genes, and especially the 5' splice site) stimulates exonic microRNA processing. Previous published data has shown a positive role of splicing and the 5' splice site in the case of animal miRNA biogenesis. However, these examples included intronic miRNAs embedded within protein-coding genes [Janas *et al.* 2011, Yan *et al.* 2012]. Janas and his colleagues showed that 5' splice site recognition by spliceosome promotes the recruitment of Drosha and microprocessing of miR-211. Moreover, a knockdown of U1 snRNP components reduced intronic miRNA expression [Janas *et al.* 2011]. On the other hand, Yan and his colleagues described the processing of a plant intronic *MIR400* transcript. They showed that the processing of miRNA400 depends on how efficiently the intron is spliced out of the host mRNA [Yan *et al.* 2012]. Together with our data, all of these experiments support the conclusion that there is an interplay between the microprocessor and splicing machinery affecting microRNA biogenesis.

5'-splice-site recognition and polyadenylation-site selection in pri-miRNA

It is noteworthy that Kurihara & Watanabe observed two types of pri-miRNA163 derived from the same gene in *Arabidopsis thaliana*. They called them type 1 and type 2 [Kurihara & Watanabe, 2004]. Both types are polyadenylated, but the polyadenylation signal of the type 1 precursor is located in the intron (giving rise to a shorter precursor), and type 2 pri-miRNA163 arise by splicing the intron and having a polyA signal in the last exon (giving rise to a longer precursor) [Kurihara & Watanabe, 2004]. Additional studies on plant pri-miRNA structures revealed that the presence of multiple polyadenylation sites in pri-miRNAs is not a marginal phenomenon [Szarzyńska *et al.* 2009, Song *et al.* 2007]. The presence of two polyadenylation signals (PASs) in the pri-miR163 transcripts prompted us to investigate an additional function of U1 snRNP, independent of its function in splicing. This additional function of U1 snRNP is the protection of transcripts from premature cleavage and polyadenylation (PCPA). Kaida and his colleagues showed that when the binding between U1 snRNA and 5' splice sites is blocked in HeLa cells, cryptic polyadenylation signals were activated [Kaida *et al.* 2010]. These cryptic PASs typically are within introns, and their activation results in the failure in production of full-length pre-mRNAs from the majority of genes. We calculated the levels of type 1 and type 2 miR163 transcripts (terminated within the intron and spliced) in wild-type *A. thaliana* plants. The

results revealed that 60% of pri-miR163 is terminated at the distal polyA site. However, in the mutants in which the 5′ splice site of the *MIR163* gene was mutated, the proximal polyadenylation site was almost exclusively selected. We suggest that the lack of U1 snRNP binding in pri-miR163 favors proximal, intronic polyA signal recognition, while the distal polyA signal is preferentially selected in wild-type pri-miRNA. Parallely, splicing inhibition and proximal polyA site selection correlated with the downregulation of exonic miR163 biogenesis. It is likely that the spliceosome assembly (aided by specific RNA-binding proteins) may control the length and structure of *MIR* gene transcripts and regulate mature miRNA levels, depending on the tissue and stage of developmental.

Crosstalk model between splicing and dicing machineries

Our work (as well as Schwab's) clearly shows that plant introns enhance miRNA production (in the case when introns localize 3' to the miRNA stem-loop structure). Based on results described in this PhD thesis, we proposed a model showing a splicing-machinery-stimulatory effect on miRNA biogenesis. A model is presented in Figure 46. The CBC complex is bound to the cap structure of all nascent RNA POL II transcripts as well as to the transcripts of *MIR* genes. CBP20 and CBP80 proteins that form the CBC complex interact with SERRATE. The SE interacts with a DCL1 protein that is an RNAse type III enzyme. We postulate that the SE can also interact with some proteins from U1 snRNP that, according to this model, will link the splicing of a downstream intron to miRNA biogenesis machinery (new data obtained in the Department of Gene Expression confirms this assumption). We showed that the splicing of an intron (and especially a functional 5` splice site) have a boosting effect on miRNA biogenesis. Additionally, the 5` splice site suppresses intronic (proximal) polyA site activation.

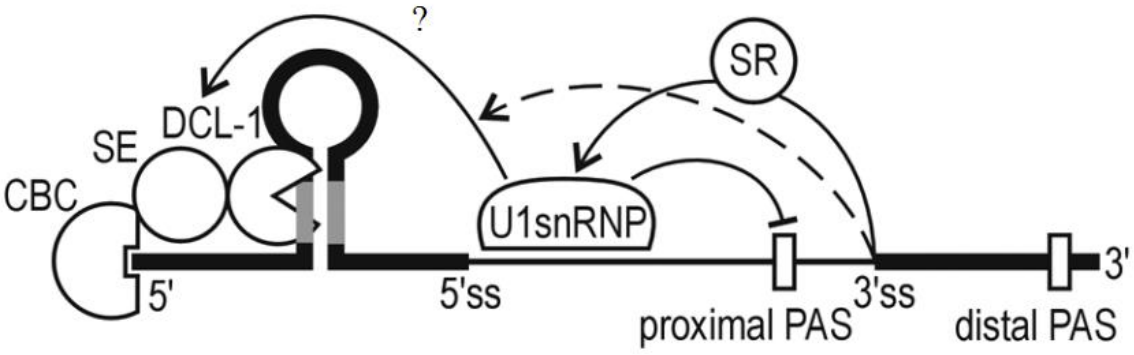


Figure 46 Model showing splicing machinery stimulatory effect on miRNA biogenesis. Thick lines represent exons, and thin lines introns. Grey parts of the stem-loop structure mark miRNA and miRNA*. Arrows show stimulatory effects, while the arrow with no head points to the inhibitory effect. Black arrows depict strong positive effects, while dashed arrows show weak stimulation. Open boxes inserted within the pri-miRNA structure represent proximal and distal polyadenylation sites (PAS). CBC, Cap-Binding Complex; SE, SERRATE; DCL1, Dicer Like 1; U1 snRNP, U1 small nuclear ribonucleoprotein particle; SR, Serine Arginine Rich Proteins; 5′ ss and 3′ ss mean the 5′ and 3′ splicing sites, respectively [Szweykowska-Kulińska *et al.* 2013,

Splicing of intron-containing pri-miRNA 163 regulates a proper response to environmental cues

In this work, we also showed that the presence of functional splice sites in pri-miR163 is mandatory for the pathogen-triggered accumulation of miR163. It was shown by Ng and his colleagues that *MIR163* is induced by various biotic stresses. In our work, we have shown that, in wild-type plants, a *Pseudomonas syringae* infection leads to the strong induction of miR163 expression and to the consequent downregulation of target mRNA levels that encode S-adenosyl-l-methionine-dependent methyltransferase (At1g66690). However, in *Arabidopsis* transgenic plants where microRNA163 was expressed from the transgene with inactivated 5’ and 3’ splice sites, *Pseudomonas syringae* treatment did not induce miR163 expression, as its level remained low. Consequently, the level of the miR163 target mRNA (At1g66690) was upregulated. These experiments show that the processing of intron-containing pri-miRNA163 is required for proper plant-defense response.

A model in which splicing influences the fine-tuning expression of *Arabidopsis* intronic miR400 was previously reported by Kang Yan and his colleagues [Yan *et al.* 2012]. According to the authors, decreased accumulation of the miR400 is a consequence of alternative splicing of the miRNA-containing intron. Under heat treatment, an alternative 5' splice site that is downstream of the miR400 sequence is selected. This selection leads to the excision of a 100-base-pair fragment from the 3' end of the first intron while leaving miR400 in the pri-miRNA transcript that leads to a decreased efficiency of processing into mature microRNA. Additionally, transgenic plants overexpressing *MIR400* under heat-stress conditions have a lower germination rate and decreased elongation of the hypocotyl when compared with wild-type plants. The obtained results demonstrate that the decreased level of miR400 under heat stress made the plants less sensitive to heat [Yan *et al.* 2012].

Another example of fine-tuning microRNA expression by splicing was described by Fan Jia & Christopher D. Rock. [Jia & Rock 2013]. MiR846 and miR842 are encoded within one transcriptional unit, but they are generated from alternatively spliced isoforms. Both miRNAs are functionally related, and their levels are regulated by abscisic acid (ABA). Both miRNAs are expressed in the roots; but upon ABA, the level of miR846 is decreased. This downregulation is accompanied by a decreased level of alternatively-spliced isoforms. All of these results strengthen the idea of the importance of posttranscriptional pri-miRNA regulation in plant life.

2 Transcriptional regulation of biogenesis of microRNAs

The HYL1 protein plays a major role in the processing of pri-miRNAs. But it was also reported that, without the HYL1 protein, both forms of pri-miRNA (before and after splicing) accumulate to high levels in *Arabidopsis thaliana* [Szarzyńska *et al*. 2009]. Therefore, it seems that HYL1 may have an additional function in microRNA biogenesis. It appears that the recruitment of HYL1 to the miRNA biogenesis machinery takes place at the very early stages of pri-miRNA processing, possibly before splicing occurs. It has been shown that, for protein-coding genes, there is a direct connection between gene transcription and further, co- and post-transcriptional processing of nascent pre-mRNAs. It is known that all (or almost all) plant *MIR* genes are transcribed by RNA Pol II. In addition, it has been found that, similar to pre-mRNAs, primary transcripts of MIR genes undergo further processing, including cap structure formation, polyadenylation, and

splicing. Moreover, the phosphorylation status of the CTD domain of RNA POL II is very important in transcription. Interplay between kinases and phosphatases acting on RNA Pol II can modify gene expression. One of the proteins that is able to dephosphorylate the CTD of the RNA POL II at serine 5 residue is a protein phosphatase called CPL1 [Koiwa *et al.* 2004]. Recently, it was reported that the phosphorylation status of HYL1 is also important in microRNA biogenesis. It was shown that the HYL1 protein needs to be dephosphorylated for its optimal activity [Manavella *et al.* 2012]. Moreover, hypophosphorylation of HYL1 is maintained by the CPL1 protein. These observations strengthen the hypothesis that HYL1 is involved in the transcription of *MIR* genes.

HYL1 influences the transcription of a GUS reporter

Our results presented in this manuscript have shown that a GUS reporter that was under control of an *MIR* gene promoter is expressed at a lower level when HYL1 is missing. On the other hand, no such effect was visible when the GUS reporter was under control of the protein-coding gene promoter. These experiments suggest that HYL1 is a necessary factor in the transcription of *MIR* genes. However, our observations may result from some indirect effects. In the *hyl1*-2 background, the majority of microRNAs are expressed at a low level, so one can imagine that the absence of a particular microRNA can cause the accumulation of a target gene that controls the transcription of a GUS reporter that is under control of a given protein-coding gene promoter. Schematic representation of a possible feedback loop is presented in Figure 47.

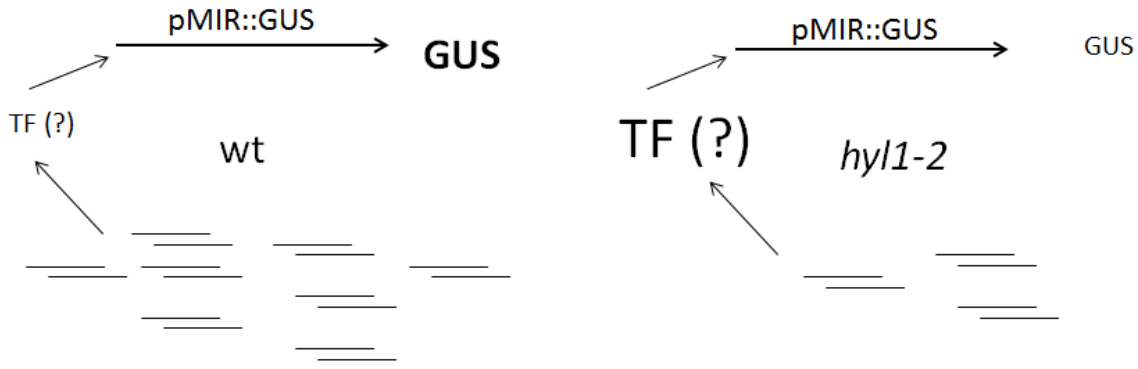


Figure 47 Schematic representation of a possible feedback loop in the *hyl1-2* mutant. In the *hyl1-2* mutant, microRNAs are far less abundant as compared to wild-type plants. The decreased level of a particular microRNA can cause the accumulation of as-of-yet unidentified transcription factor(s). This transcription factor might be involved in the transcription of a reporter gene that is under the promoter that was used in our studies, which in the end can result in lower expression of GUS transcripts and, as a consequence, the GUS protein. TF, Transcription factor; wt, wild type.

To rule out such a possibility, we decided to cross transgenic lines with GUS reporters that were under different promoters into a selected microRNA biogenesis mutant background. In this microRNA biogenesis mutant, the majority of microRNAs should be expressed at a low level, but the protein that is mutated should not be involved in the early stages of microRNAs biogenesis. A good candidate is a *hen1-5* mutant. The HEN1 protein recognizes 2-nt-long 3' overhangs of miRNA/miRNA* duplex and catalyzes 2`-O-methylation of the last 3` nucleotides, which increases their stability [Yang *et al.* 2006b]. In the *hen1-5* mutant, the majority of microRNAs are expressed at a low level, and phenotypically, this mutant resembles the phenotype of *hyl1-2* to some extent. If a GUS reporter that is under control of a *MIR* gene promoter would be at the same level in wild-type plants and the *hen1-5* mutant, it will mean that no negative loop between the GUS reporter and the low level of microRNAs exist in *Arabidopsis thaliana.* Crosses between GUS reporter lines and the *hen1-5* mutant are done, and preliminary data showed that the level of the GUS protein is at the same level in the tested backgrounds. However, staining was done only in the case of two plants; so, to be sure that no negative loop exists, we have to confirm the results of GUS staining in more plants.

HYL1 influences the RNA POL II distribution

Evaluation of the occupancy of RNA Pol II on *MIR* genes in *hyl1-2* showed an accumulation of total RNA Pol II in the region at approximately 200 base pair upstream of the transcription initiation site. This phenomenon was observed in the cases of *MIR393A*, *MIR393B,* and *MIR166A* in the *hyl1-2* mutant. While the distribution of RNA POL II in a variety of complex genomes is correlated with gene expression, the presence of RNA POL II at a specific locus does not necessarily indicate active expression from this locus. The higher occupancy of the RNA polymerase II at the analyzed *MIR* promoter regions, together with results from the GUS reporter assays, suggest that transcription transition from initiation to elongation is downregulated in the *hyl1-2* mutant. A model showing HYL1's influence on *MIR* gene transcription is shown in Figure 48.

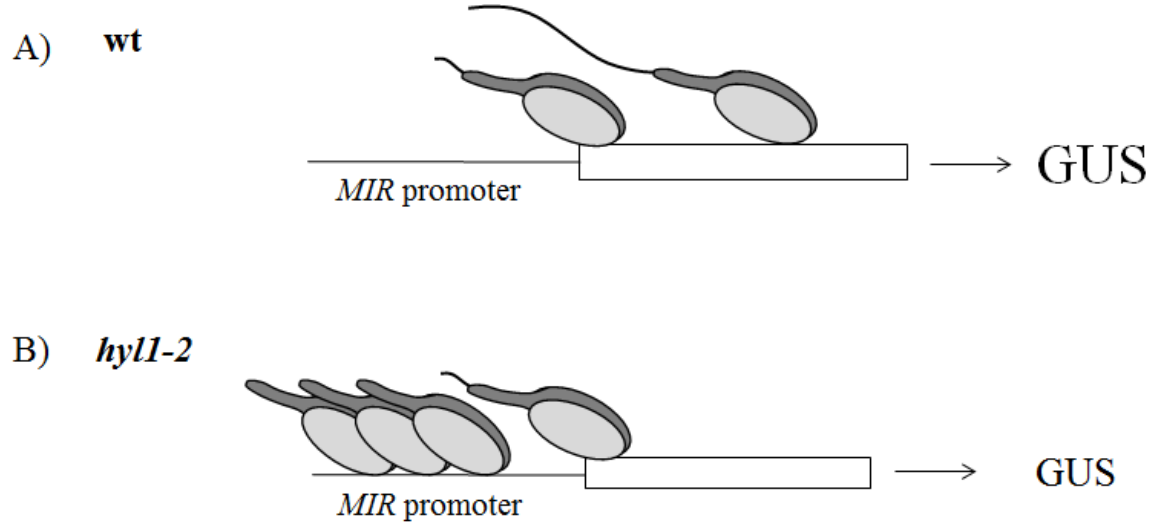


Figure 48 Schematic representation of a model showing HYL1's influence on a GUS reporter. In the *hyl1-2* mutant (B) as compared to wild-type plants (A), RNA POL II accumulates at the region upstream of the transcription initiation site. Accumulation of RNA POL II leads to lower expression of the GUS transcript.

Mutation of the CPL1 protein also leads to an accumulation of total RNA pol II at the region upstream of the TSS, and this effect was visible in all of the tested *MIR* genes (please compare Figures 33 and 34C). Pablo Manavella and his colleagues showed that pri-miRNAs and their corresponding microRNAs are downregulated in the *cpl1-7* mutant as compared to wild-type plants [Manavella *et al.* 2012]. These observations indirectly suggest that the accumulation of RNA POL II does not indicate active expression. However, we can only speculate due to the lack of appropriate GUS reporter lines crossed into the *cpl1-7* background.

***MIR* gene promoter activity is sensitive to the HYL1 protein**

The obtained results showed that replacement of the native promoter of the *MIR163* gene with a CaMV 35S promoter (or other protein-coding gene promoters) makes the transcription of pri-miRNA 163 insensitive to HYL1 proteins. Pri-microRNA163 that is transcribed from other than native promoters is at the same level in wild-type and *hyl1-2* plants (results are presented in Figure 36). Irrespective of which promoter was used to control *MIR163* gene expression, the level of mature microRNA 163 was decreased in the hyl1-2 mutant in all cases. HYL1, as a partner of DCL1, is mainly involved in the processing of pri-miRNA through pre-miRNA to mature microRNA. Using Northern-blot hybridization, we observed mostly a consequence of the lack of HYL1 in the processing of pri-miRNA. The best way to discriminate between the processing and transcription of *MIR*

genes in the context of HYL1 would be to make experiments on a mutant that lacks the HYL1 protein, but the processing of pri-miRNAs is not affected. A dominant mutation in DCL1 (*dcl1-13*) that causes amino acid substitution in the RNA helicase domain confers significant restoration in the level of miRNAs in the *hyl1-2* mutant [Tagami *et al.* 2009]. Transformation of double *dcl1-13/hyl1-2* with constructs of *MIR163* under control of different promoters should let us observe the influence of HYL1 on only the transcription, not on the processing of pri-miRNAs.

PERSPECTIVES

1. **Analysis of the level of pri-miRNA and mature microRNA in U1 snRNP mutants**

   The presented experiments revealed a connection between the dicing and splicing of plant pri-miRNAs. We observed that splicing machinery stimulates miRNA biogenesis. Moreover, the boosting effect in the miRNA biogenesis depends on a functional 5` splice that is recognized by U1 snRNP. It would be interesting to investigate microRNA processing in the mutants of some components of U1 snRNP. A decreased level of U1 snRNP in cells should influence the level of mature microRNA. We also have to take into consideration the potential indirect effects that can be caused by a wide range of splicing events. Experiments relevant to this issue will be performed in the Department of Gene Expression.

2. **Studies on the interaction between the SERRATE protein and components from U1 snRNP**

   U1snRNP is a multiprotein large nuclear complex. It is possible that some of the components of U1 snRNP interact with a major protein player from the microRNA biogenesis machinery. SERRATE protein is a good candidate for this interaction. SE is a large protein that is involved in the splicing of pre-mRNAs. SERRATE mostly influences the selection of the 5` alternative splice site in pre-mRNAs. Interaction between SERRATE and components of U1 snRNP will be tested in the Department of Gene Expression by a yeast-two hybrid system and by fluorescence microscopy assays.

3. **Analysis of polyA site usage and its involvement in the maturation of microRNAs**

   Mutation of the 5' splice site in the *MIR163* gene interferes with polyA site selection within pri-miRNA163. As a consequence, production of microRNA from such transcripts is less efficient. It would be interesting to discriminate the effect on microRNA processing between splicing and polyA usage. One of the possible experiments would be to prepare genetic constructs where the native intron of *MIR163* is replaced by a similar-sized intron having no polyA signal. Analysis of the processing of microRNA after additional mutation of the 5' splice site in this new construct could help us better understand the interplay between splicing and

polyA machinery. Analysis of microRNA levels in polyA polymerase mutants would be another step in such experiments. Experiments relevant to this issue will be performed in the Department of Gene Expression.

4. **Analysis of a connection between the dicing and splicing in an intronic MIR gene**

   In the case of independent transcription units of *MIR* genes, we showed a connection between the dicing and splicing of plant pri-miRNAs. It would be interested to test whether the mutation of six nucleotides within the 5' splice site in an intronic *MIR* gene would affect the production of microRNA. Moreover, it would also be interesting to test the influence of splice-site mutation on polyadenylation usage within the microRNA precursor. And finally, what might happen with the maturation of microRNA after changing the location of the stem-loop structure from intronic to exonic? Experiments relevant to this issue will be performed in the Department of Gene Expression.

5. **Analysis in a global view of POL II distribution in a *hyl1-2* mutant**

   Chromatin Immunoprecipitation followed by high-throughput sequencing is a valuable method for studying interactions between proteins and DNA. Immunoprecipitation of chromatin will be done using antibodies against total RNA POL II in wild-type plants and the *hyl1-2* mutant. A high-resolution map of POL II occupancy across the entire *Arabidopsis thaliana* genome in *hyl1-2* will allow us to observe at which *loci* HYL1 mostly affects RNA POL II accumulation. Experiments relevant to this issue will be performed in the Department of Gene Expression.

6. **Analysis of pri-miRNA and mature microfigurRNA levels in double *cpl1-7 hyl1-2* mutants**

   In a single *hyl1-2* and *cpl1-7* mutant, RNA POL II accumulated at a region upstream of the TSS. Pri-microRNAs accumulate in *hyl1-2* as a consequence of ineffective processing, and in *cpl1-7*, pri-microRNAs also accumulate (but probably as a consequence of inefficient transcription). The level of microRNAs in both mutants is severely decreased. Decreased accumulation of microRNAs was more prominent in the *hyl1-2* mutant than in the *cpl1-7* because, in *Arabidopsis*

thaliana, there is also a CPL2 protein that is a close homolog of CPL1 and may overtake its function. The double mutant of CPL1 and CPL2 is embryonically lethal [Kowia *et al.* 2004]. It would be interesting to make a double mutant of *hyl1-2* and *cpl1-7*, test RNA POL II occupation at the *MIR* gene promoter, and also evaluate the levels of pri-miRNAs and mature microRNAs.

7. Analysis of proteins interacting with HYL1 by mass spectroscopy

HYL is involved in the early stages of the biogenesis of microRNAs in *Arabidopsis thaliana*. It was reported that the HYL1 protein interacts with components of the DCL1 complex, but also with the CPL1 protein that interacts with RNA POL II. It would be interesting to know other proteins from the transcription complex that interact with the HYL1 protein. To explore these new interacting proteins, HYL1 (under its native promoter) will be expressed with a tag in *Arabidopsis thaliana*. Plant tissue before immunoprecipitation will be crosslinked with formaldehyde, which will also help us find weak protein-protein interactions. Immunoprecipitation will be followed by mass spectroscopy analysis. Moreover, the interaction between HYL1 and good candidates from mass-spectroscopy analysis should be confirmed using other techniques; for example, the FLIM-FRET, pull-down, or yeast-two hybrid systems.

LITERATURE

Allen E, Xie Z, Gustafson AM, Carrington JC (2005) microRNA-directed phasing during *trans*-acting siRNA biogenesis in plants. *Cell* **121**, 207-21.

Aukerman MJ, Sakai H (2003) Regulation of flowering time and floral organ identity by a microRNA and its APETALA-like target genes. *Plant Cell* **15**, 2730-41.

Baek D, Park HC, Kim MC, Yun DJ (2013) The role of Arabidopsis MYB2 in miR399f-mediated phosphate-starvation response. *Plant Signal Behav* **8**, e23488.

Barabino SM, Blencowe BJ, Ryder U, Sproat BS, Lamond AI (1990) Targeted snRNP depletion reveals an additional role for mammalian U1 snRNP in spliceosome assembly. *Cell* **63**, 293–302.

Berget SM, Moore C, Sharp PA (1977) Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *Proc Natl Acad Sci USA* **74**, 3171-5.

Beyer AL, Osheim YN (1988) Splice site selection, rate of splicing, and alternative splicing on nascent transcripts. *Genes Dev* **2**, 754-65.

Bielewicz D, Kalak M, Kalyna M, Windels D, Barta A, Vazquez F, Szweykowska-Kulińska Z, Jarmołowski A (2013) Introns of plant pri-miRNAs enhance miRNA biogenesis. *EMBO Rep* **14**, 622-8.

Bollman KM, Aukerman MJ, Park MY, Hunter C, Berardini TZ, Poethig RS (2003) HAST, the *Arabidopsis* ortholog of exportin 5/MSN5, regulates phase change and morphogenesis. *Development* **130**, 1493-504.

Bologna NG, Mateos JL, Bresso EG, Palatnik JF (2009) A loop-to-base processing mechanism underlies the biogenesis of plant microRNAs miR319 and miR159. *EMBO J* **28**, 3646-56.

Boutet S, Vazquez F, Liu J, Beclin C, Fagard M, Gratias A, Morel JB, Crete P, Chen X, Vaucheret H (2003) Arabidopsis HEN1: a genetic link between endogenous miRNA controlling development and siRNA controlling transgenic silencing and virus resistance. *Curr Biol* **13**, 843-8.

Brown JWS, Marshall DF, Echeverria M (2008) Intronic noncoding RNAs and splicing. *Trends Plant Sci* **13**, 335-42.

Cho SK, Chaabane SB, Shah P, Poulsen CP, Yang SW (2014) COP1 E3 ligase protects HYL1 to retain microRNA biogenesis. *Nat Commun* **5,**5867.

Clough SJ, Bent AF (1998) *Floral dip*: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J* **16**, 735-43.

Company M, Arenas J, Abelson J (1991) Requirement of the RNA helicase-like protein PRP22 for release of messenger RNA from spliceosomes. *Nature* **349**, 487–93.

Curtis MD, Grossniklaus U (2003) A gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant Physiol* **133**, 462-9.

Gruber JJ, Zatechka DS, Sabin LR, Yong J, Lum JJ, Kong M, Zong WX, Zhang Z, Lau CK, Rawlings J, Cherry S, Ihle JN, Dreyfuss G, Thompson CB (2009) Ars2 links the nuclear cap-binging complex to RNA interference and cell proliferation. *Cell* **138**, 328-39.

Dong Z, Han MH, Fedoroff N (2008) The RNA-binding proteins HYL1 and SE promote *in vitro* processing of pri-miRNA by DCL1. *Proc Natl Acad Sci USA* **105**, 9851-52.

Eamens AL, Smith NA, Curtin SJ, Wang MB, Waterhouse PM (2009) The *Arabidopsis thaliana* double-stranded RNA binding protein DRB1 directs guide strand selection from microRNA duplexes. *RNA* **15**, 2219-35.

Gasciolli V, Mallory AC, Bartel DP, Vaucheret H (2005) Partially redundant functions of *Arabidopsis* DICER-like enzymes and a role for DCL4 in producing *trans*-acting siRNAs. *Curr Biol* **15**, 1494-1500.

Golovkin M, Reddy AS (1998) The plant U1 small nuclear ribonucleoprotein particle 70K protein interacts with two novel serine/arginine-rich proteins. *Plant Cell* **10**, 1637-48.

Gregory BD, O'Malley RC, Lister R, Urich MA, Tonti-Filippini J, Chen H, Millar AH, Ecker JR (2008) A link between RNA metabolism and silencing affecting *Arabidopsis* development. *Dev Cell* **14**, 1-13.

Grigg SP, Canales C, Hay A, Tsiantis M (2005) SERRATE coordinates shoot meristem function and leaf axial patterning in *Arabidopsis. Nature* **437**, 1022-26.

Hirsch J, Lefort V, Vankersschaver M, Boualem A, Lucas A, Thermes C, d'Aubenton-Carafa Y, Crespi M (2006) Characterisation of a 43 non-protein-coding mRNA genes in *Arabidopsis*, including the *MIR162a*-derived transcripts. *Plant Physiol* **140**, 1192-1204.

Hafner M, Lianoglou S, Tuschl T, Betel D (2012) Genome-wide identification of miRNA targets by PAR-CLIP. *Methods* **58**, 94-105

Hirano SS, Upper CD (2000) Bacteria in the leaf ecosystem with emphasis on *Pseudomonas syringae-a* pathogen, ice nucleus, and epiphyte. *Microbiol Mol Biol Rev* **64**, 624-53.

Huang Y, Ji L, Huang Q, Vassylyvev DG, Chen X, Ma JB (2009) Structural insights into mechanisms of the small RNA methyltransferase HEN1. *Nature* **461**, 823-7.

Iki T, Yoshikawa M, Meshi T, Ishikawa M (2012) Cyclophilin 40 facilitates HSP90-mediated RISC assembly in plants. *EMBO J* **31**, 267-78.

Iki T, Yoshikawa M, Nishikiori M, Jaudal MC, Matsumoto-Yokoyama E, Mitsuhara I, Ishikawa M (2010) In vitro assembly of plant RNA-induced silencing complexes facilitated by molecular chaperone HSP90. *Molecular cell* **39**, 282-91.

Izaurralde E, Lewis J, McGuigan C, Jankowska M, Darzynkiewicz E, Mattaj IW (1994) A nuclear cap binding protein complex involved in pre-mRNA splicing. *Cell* **78**, 657-68.

Izaurralde E, Lewis J, Gamberi Ch, Jarmolowski A, McGuigan C, Mattaj IW (1995) A cap-binding protein complex mediating U snRNA export. *Nature* **376**, 709-12.

Jamison SF, Pasman Z, Wang J, WillC, Lührmann R, Manley JL, Garcia-Bianco MA (1995) U1 snRNP-ASF/SF2 interaction and 5` splice site recognition: characterization of required elements. *Nucleic Acids Res* **23**, 3260-7.

Janas MM, Khaled M, Schubert S, Bernstein JG, Golan D, Vequilla RA, Fisher DE, Shomron N, Levy C, Novina CD (2011) Feed-forward microprocessing and splicing activities at a microRNA-containing intron. *PLoS Genet.* **7**, e1002330

Jeong IS, Aksoy E, Fukudome A, Akhter S, Hiragauri A, Fukuhara T, Bahk JD, Koiwa H (2013) Arabidopsis C-terminal domain phosphatase-like 1 functions in miRNA accumulation and DNA methylation. *PLoS One* **8**, e74739.

Jia F, Rock CD (2013) MIR846 and MIR842 comprise a cistronic MIRNA pair thatis regulated by abscisic acid by alternative splicing in roots of Arabidopsis. *Plant Mol Biol* **81**, 447-60.

Kaida D, Berg MG, Younis I, Kasim M, Singh LN, Wan L, Dreyfuss G (2010) U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. *Nature* **468**, 664-8.

Kandels-Lewis S, Seraphin B (1993) Involvement of U6 snRNA in 5 splice site selection. *Science* **262**, 2035–9.

Kim YJ, Zheng B, Yu Y, Won SY, Mo B, Chen X (2011) The role of Mediator in small and long noncoding RNA production in *Arabidopsis thaliana*. *EMBO J* **30**, 814-22.

Kmieciak M, Simpson CG, Lewandowska D, Brown JW, Jarmolowski A (2002) Cloning and characterization of two subunits of *Arabidopsis thaliana* nuclear cap-binding complex. *Gene* **283**, 171-83.

Koiwa H, Hausmann S, Bang WY, Ueda A, Kondo N, Hiraguri A, Fukuhara T, Bahk JD, Yun DJ, Bressan RA, Hasegawa PM, Shuman S (2004) Arabidopsis C-terminal domain phosphatase-like 1 and 2 are essential Ser-5-specific C-terminal domain phosphatases. *Proc Natl Acad Sci USA* **101**, 14539-44.

Koster T, Meyer K, Weinholdt C, Smith LM, Lummer M, Speth C, Grosse I, Weigel D, Staiger D (2014) Regulation of pri-miRNA processing by the hnRNP-like protein AtGRP7 in Arabidopsis. *Nucleic Acids Res* **41**, 9925-36.

Król J, Loedige I, Filipowicz W (2010) The widespread regulation of microRNA biogenesis, function and decay. *Nat Rev Genet* **11**, 597-610.

Kruszka K, Pieczynski M, Windels D, Bielewicz D, Jarmolowski A, Szweykowska-Kulińska Z, Vazquez F (2012) Role of microRNAs and other sRNAs of plants in their changing environments. *J Plant Phsysiol* **169**, 1664-72.

Kuhn JM, Breton G, Schroeder JI (2007) mRNA metabolism of flowering-time regulators in wild-type *Arabidopsis* revealed by a nuclear cap binding protein mutant, *abh1*. *Plant J* **50**, 1049-62.

Kurihara Y, Watanabe Y (2004) *Arabidopsis* microRNA biogenesis through Dicer-like 1 protein functions. *Proc Natl Acad Sci USA* **101**, 12753-58.

Lagos-Quintana M, Rauhut R, Lendeckel W, Tuschl T (2001) Identification of novel genes coding for small expressed RNAs. *Science* **294**, 853-8.

Lau NC, Lim LP, Weinstein EG, Bartel DP (2001) An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* **294**, 858-62.

Laubinger S, Sachsenberg T, Zeller G, Busch W, Lohmann JU, Raetsch G, Weigel D (2008) Dual roles of the nuclear cap-binding complex and SERRATE in pre-mRNA splicing and microRNA processing in *Arabidopsis thaliana*. *Proc Natl Acad Sci USA* **105**, 8795-8800.

Lee R, Feinbaum R, Ambros V (2004) A short history of a short RNA.*Cell* **116**, 89-92.

Lee RC, Ambros V (2001) An extensive class of small RNAs in *Caenorhabditis elegans*. *Science* **294**, 862-4

Lee RC, Feinbaum RL, Ambros V (1993) The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* **75**, 843-54.

Lewandowska D, Simpson CG, Clark GP, Jennings NS, Barciszewska-Pacak M, Lin CF, Makalowski W, Brown JW, Jarmolowski A (2004) Determinants of plant U12-dependent intron splicing efficiency. *Plant Cell* **16**, 1340-52.

Li S, Liu L, Zhuang X, Yu Y, Liu X, Cui X, Ji L, Pan Z, Cao X, Mo B, Zhang F, Raikhel N, Jiang L, Chen X (2013) MicroRNAs inhibit the translation of target mRNAs on the endoplasmic reticulum in Arabidopsis. *Cell* **153**, 562-74.

Lobbes D, Rallapalli G, Schmidt DD, Martin C, Clarke J (2006) SERRATE: a new player on the plant microRNA scene. *EMBO Rep* 7, 1052-58.

Machida S, Yuan YA (2013) Crystal structure of Arabidopsis thaliana Dawdle forkhead-associated domain reveals a conserved phospho-threonine recognition cleft for dicer-like 1 binding. *Mol Plant* **6**, 1290-300.

Mallory AC, Reinhart BJ, Jones-Rhoades MW, Tang G, Zamore PD, Barton MK, Bartel DP (2004) MicroRNA control of PHABULOSA in leaf development: importance of pairing to the microRNA 5' region. *EMBO J* **23**, 3356-64.

Manavella PA, Hagmann J, Ott F, Laubinger S, Franz M, Macek B, Weigel D (2012) Fast-forward genetics identifies plant CPL phosphatases as regulators of miRNA processing factor HYL1. *Cell* 151, 859-70.

Morris ER, Chevalier D, Walker JC (2006) *DAWDLE*, a forkhead-associated domain gene, regulates multiple aspects of plant development. *Plant Physiol* **141**, 932–41.

Neugebauer KM (2002) On the importance of being co-transcriptional. *J Cell Sci* **115**, 3865-71.

Ng DW, Zhang C, Miller M, Palmer G, Whiteley M, Tholl D, Chen ZJ (2011) cis- and trans- Regulation of miR163 and target genes confers natural variation of secondary metabolites in two *Arabidopsis* species and their allopolyploids. *Plant Cell* **23**, 1729-40.

Nicovics K, Blein T, Peaucelle A, Ishida T, Morin H, Aida M, Laufs P (2006) The balance between the *MIR164A* and *CUC2* genes controls leaf margin serration in *Arabidopsis*. *Plant Cell* **18**, 2929-45.

Padgett RA, Grabowski PJ, Konarska MM, Seiler S, Sharp PA (1986) Splicing of messenger RNA precursors. *Annu Rev Biochem* **55**, 1119-50.

Papp I, Mur LA, Dalmadi A, Dulai S, Koncz C (2004) A mutation in the Cap Binging Protein 20 gene confers frought tolerance to *Arabidopsis*. *Plant Mol Biol* **55**, 679-86.

Park MY, Wu G, Gonzalez-Sulser A, Vaucheret H, Poethig RS (2005) Nuclear processing and export of microRNAs in *Arabidopsis*. *Proc Natl Acad Sci USA* **102**, 3691-6.

Park W, Li J, Song R, Messing J, Chen X (2002) CARPEL FACTORY, a Dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in *Arabidopsis thaliana. Curr Biol* **12**, 1484-95.

Parker JS, Roe SM, Barford D (2004) Crystal structure of a PIWI protein suggests mechanisms for siRNA recognition andslicer activity. *EMBO J* **23**, 4727-37.

Parry G, Calderon-Vilalobos LI, Prigge M, Peret B, Dharmasiri S, Itoh H, Lechner E, Gray WM, Bennet M, Estelle M (2009) Complex regulation of the TIR1/AFB family of auxin receptors. *Proc Natl Acad Sci USA* **106**, 22540-5.

Raczyńska KD, Simpson CG, Ciesiolka A, Szewc L, Lewandowska D, McNicol J, Szweykowska-Kulińska, Brown JW, Jarmolowski A (2010) Involvement of the nuclear cap-binding protein complex in alternative splicing in *Arabidopsis thaliana*. *Nucleic Acids Res* **38**, 265-78.

Raczyńska KD, Stepien A, Kierzkowski D, Kalak M, Bajczyk M, McNicol J, Simpson CG, Szweykowska-Kulińska Z, Brown JW, Jarmolowski A (2012) The SERRATE protein is involved in alternative splicing in *Arabidopsis thaliana*. *Nucleic Acids Res* **42**, 1224-44.

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

Reddy AS, Marquez Y, Kalyna M, Barta A (2013) Complexity of the alternative splicing landscape in plants. *Plant Cell* **25**, 3657-83.

Reddy ASN (2007) Alternative splicing of pre-messenger RNAs in plants in the genomic era. *Annu Rev Plant Biol* **58**, 267-294.

Reed R (1996) Initial splice-site recognition and pairing during pre-mRNA splicing. *Curr Opin Genet Dev* **6**, 215–20.

Reinhart BJ, Slack FJ, Basson M, Pasquielli AE, Bettinger JC, Rougvie AE, Horvitz HR, Ruvkun G (2000) The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature* **403**, 901-6.

Reis RS, Hart-Smith G, Eamens AL, Wilkins MR, Waterhouse PM (2015) Gene regulation by translational inhibition is determined by Dicer partnerning proteins. *Nature plants* **1**, 14027

Ren G, Xie M, Dou Y, Zhang S, Zhang C, Yu B (2012a) Regulation of miRNA abundance by RNA binding protein TOUGH in Arabidopsis. *Proc Natl Acad Sci USA* **109**, 12817-21.

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

Rivas FV, Tolia NH, Song JJ, Aragon JP, Liu J, Hannon GJ, Joshua-Tor L (2005) Purified Argonaute2 and an siRNA form recombinant human RISC. *Nat Struct Mol Biol* **12**, 340-9.

Rodriguez A, Griffiths-Jones S, Ashurst JL, Bradley A (2004) Identification of mammalian microRNA host genes and transcription units. *Genome Res* **14**, 1902-10.

Rogers K, Chen X (2013) Biogenesis, Turnover, and Mode of Action of Plant MicroRNAs. *Plant Cell* **25**, 2383-99.

Roy J, Kim K, Maddock JR, Anthony JG, Woolford JL Jr (1995) The final stages of spliceosome maturation require Spp2p that can interact with the DEAH box protein Prp2p and promote step 1 of splicing. *RNA* **1**, 375–390.

Sabin LR, Zhou R, Gruber JJ, Lukinova N, Bambina S, Berman A, Lau CK, Thompson CB, Cherry S (2009) Ars2 regulates both miRNA- and siRNA- dependent silencing and suppresses RNA virus infection in *Drosophila*. *Cell* **138**, 340-51.

Schwab R, Speth C, Laubinger S, Voinnet O (2013) Enhanced microRNA accumulation through stemloop-adjacent introns. *EMBO Rep* **14**, 615-21.

Schwer B, Gross CH (1998) Prp22, a DExH-box RNA helicase, plays two distinct roles in yeast pre-mRNA splicing. *EMBO J* **17**, 2086–94.

Sharp PA (2005) The discovery of split genes and RNA splicing. *Trends Biochem Sci* **30**, 279-81.

Sobkowiak L, Karłowski W, Jarmołowski A, Szweykowska-Kulińska Z (2012) Non-Canonical Processing of *Arabidopsis* pri-miR319a/b/c Generates Additional microRNAs to Target One RAP2.12 mRNA Isoform. *Front Plant Sci* **19**, 3:46.

Song L, Han M-H, Lesicka J, Fedoroff N (2007) *Arabidopsis* primary microRNA processing proteins HYL1 and DCL1 define a nuclear body distinct from the Cajal body. *Proc Natl Acad Sci USA* **104**, 5437-42.

Speth C, Willing EM, Rausch S, Schneeberger K, Laubinger S (2013) RACK1 scaffold proteins influence miRNA abundance in Arabidopsis. *Plant J* **76**, 433-45.

Speth C, Laubinger S (2014) RACK1 and the microRNA pathway: is it deja-vu all over again? *Plant Signal Behav* **9**, e27909.

Szarzyńska B, Sobkowiak L, Pant DB, Balazadeh S, Scheible WR, Mueller-Roeber B, Jarmolowski A, Szweykowska-Kulińska Z (2009) Gene structures and processing of Arabidopsis thaliana HYL1-dependent pri-miRNAs. *Nucleic Acids Res* **37**, 3083-93.

Szweykowska-Kullińska Z, Jarmołowski A, Vazquez F (2013) The crosstalk between plant microRNA biogenesis factors and the spliceosome. *Plant Signal Behav* **8**, e26955.

Tagami Y, Motose H, Watanabe Y (2009) A dominant mutation in DCL1 suppresses the hyl1 mutant phenotype by promoting the processing of miRNA. *RNA* **15**, 450-8.

Tarn WY, Steitz JA (1997) Pre-mRNA splicing: the discovery of a new spliceosome doubles the challenge. *Trends Biochem Sci* **22**, 32–7.

Teigelkamp S, McGarvey M, Plumpton M, Beggs JD (1994) The splicing factor PRP2, a putative RNA helicase, interacts directly with pre-mRNA. *EMBO J* **13**, 888–97.

Vazquez F, Gasciolli V, Crete P, Vaucheret H (2004a) The nuclear dsRNA binding protein HYL1 is required for microRNA accumulation and plant development, but not posttranscriptional transgene silencing. *Curr Biol* **14**, 346-51.

Vazquez F, Vaucheret H, Rajagopalan R, Pepers C, Gasciolli V, Mallory AC, Hilbert JL, Bartel DP, Crete P (2004b) Endogenous trans-acting siRNAs regulate the accumulation of *Arabidopsis* miRNAs. *Moll Cell* **16**, 69-79.

Vazquez F (2006) *Arabidopsis* endogenous small RNAs: highways and bywas. *Trends Plant Sci* **11**, 460-8.

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

Waibel F, Filipowicz W (1990) U6 snRNA genes of Arabidopsis are transcribed by RNA polymerase III but contain the same two upstream promoter elements as RNA polymerase II-transcribed U-snRNA genes. *Nucleic Acids Res* **18**, 3451–58.

Wang BB, Brendel V (2004) The ASRG database: identification and survey of *Arabidopsis thaliana* genes involved in pre-mRNA splicing. *Genome Biol* **5**, R102.

Wang L, Song X, Gu L, Li X, Cao S, Chu C, Cui X, Chen X, Cao X (2013) NOT2 proteins promote polymerase II-dependent transcription and interact with multiple MicroRNA biogenesis factors in Arabidopsis. *Plant Cell* **25**, 715-27.

Wierzbicki AT, Haag JR, Pikaard CS (2008) Noncoding transcription by RNA polymerase Pol IVb/Pol V mediates transcriptional silencing of overlapping and adjacent genes. *Cell* **135**, 635-48.

Will CL, Lührmann R (2001) Spliceosomal UsnRNP biogenesis, structure and function. *Curr Opin Cell Biol* **13**, 290–301.

Wu X, Shi Y, Li J, Xu L, Fang Y, Li X, Qi Y (2013) A role for the RNA-binding protein MOS2 in microRNA maturation in Arabidopsis. *Cell Res* **23**, 645-57.

Xie Z, Allen E, Fahlgren N, Calamar A, Givan SA, Carrington JC (2005b) Expression of *Arabidopsis MIRNA* genes. *Plant Physiol* **138**, 2145-54.

Xie Z, Johansen LK, Gustafson AM, Kasschau KD, Lellis AD, Zilberman D, Jacobsen SE, Carrington JC (2004) Genetic and functional diversification of small RNA pathways in plants. *PLoS Biol* **2**, 0642-52.

Xie, Z, Allen E, Wilken A, Carrington JC (2005a). DICER-LIKE 4 functions in trans-acting small interfering RNA biogenesis and vegetative phase change in *Arabidopsis thaliana*. *Proc Natl Acad Sci USA* **102**, 12984-9.

Xie Z, Allen E, Fahlgren, CalamarA, Givan SA, Carrington JC (2005b) Expression of *Arabidopsis* MIRNA genes. *Plant Physiol* **138**, 2145-54.

Yamasaki H, Hayashi M, Fukazawa M, Kobayashi Y, Shikanai T (2009) SQUAMOSA promoter binding protein–like7 is a central regulator for copper homeostasis in Arabidopsis. *Plant Cell* **21**, 347-61.

Yan K, Liu P, Wu CA, Yang GD, Xu R, Guo QH, Huang JG, Zheng CC (2012) Stress-induced alternative splicing provides a mechanism for the regulation of microRNA processing in Arabidopsis thaliana. *Mol Cell* **48**, 521-31.

Yang L, Liu Z, Lu F, Dong A, Huang H (2006a) SERRATE is a novel nuclear regulator in primary microRNA processing in *Arabidopsis*. *Plant J* **47**, 841-50.

Yang Z, Ebriht YW, Yu B, Chen X (2006b) HEN1 recognizes 21-24 nt small RNA duplexes and deposits a methyl group onto the 2'OH of the 3' terminal nucleotide. *Nucl Acids Res* **34**, 667-75.

Yant L, Mathieu J, Dinh TT, Ott F, Lanz C, Wollmann H, Schmid M (2010) Orchestration of the floral transition and floral development in Arabidopsis by the bifunctional transcription factor APETALA2. *Plant Cell* **22**, 2156-70.

Yu B, Bi L, Zheng B, Ji L, Chevalier D, Agarwal M, Ramachandran V, Li W, Lagrange T, Walker JC, Chen X (2008) The FHA domain proteins DAWDLE in *Arabidopsis* and SNIP1 in humans act in small RNA biogenesis. *Proc Natl Acad Sci USA* **105**, 10073-78.

Zhan X, Wang B, Li H, Liu R, Kalia RK, Zhu JK, Chinnusamy V (2012) Arabidopsis proline-rich protein important for development and abiotic stress tolerance is involved in microRNA biogenesis. *Proc Natl Acad Sci USA* **109**, 19198-203.

Zhang S, Liu Y, Yu B (2014) PRL1, an RNA-Binding Protein, Positively Regulates the Accumulation of miRNAs and siRNAs in Arabidopsis. *PLoS Genet* **10**, e:1004841.

Zhang S, Xie M, Ren G, Yu B (2013) CDC5, a DNA binding protein, positively regulates posttranscriptional processing and/or transcription of primary microRNA transcripts. *Proc Natl Acad Sci USA* **110**, 17588-93.