

Universidade de Lisboa
Faculdade de Ciências
Departamento de Biologia Vegetal



**The feedback regulation of miRNA biogenesis in
response to water deficit in *Medicago truncatula***

Cláudio Francisco Soares Capitão

Mestrado em Biologia Celular e Biotecnologia

2009

Universidade de Lisboa
Faculdade de Ciências
Departamento de Biologia Vegetal



**The feedback regulation of miRNA biogenesis in
response to water deficit in *Medicago truncatula***

Cláudio Francisco Soares Capitão

Tese orientada por:

Professor Doutor Manuel Pedro Salema Fevereiro

(Orientador interno)

Doutora Dulce Maria Metelo Fernandes dos Santos

(Orientador externo)

Mestrado em Biologia Celular e Biotecnologia

2009

Index

1. Resumo da tese.....	3
2. Resumo	6
3. Abstract.....	7
4. Introduction.....	8
4.1. <i>Medicago truncatula</i> as a model legume.....	8
4.2. Plant Responses to Water Deficit	8
4.3. What are miRNAs?.....	9
4.4. miRNA Biogenesis.....	10
4.5. Mode of action of Plant miRNAs	12
4.6. miRNAs discovery in Plants.....	12
4.7. miRNA function in Plants.....	13
4.8. Feedback Regulation	14
5. Objectives	15
6. Methods.....	16
6.1. Plant material and treatment conditions.....	16
6.2. RNA extraction.....	16
6.3. miRNA prediction	17
6.4. Identification of genes and their characterization.....	19
6.5. Alignment and phylogenetic trees.....	20
6.6. Quantitative Real Time PCR (qRT-PCR).....	21
6.7. Northern Blot.....	22
7. Part 1	25
7.1. Introduction	25
7.2. Results	26
7.3. Discussion	29
8. Part 2	30
8.1. Introduction	30
8.2. Results	30
8.3. Discussion	35
9. Part 3	37
9.1. Introduction	37
9.2. Results	38
9.3. Discussion	42
10. Conclusions	45
11. Future Perspectives	46
12. Bibliography.....	47
13. Supplementary Tables and Figures.....	54
13.1. Methods - Supplementary Tables and Figures	54
13.2. Part 1 - Supplementary Tables and Figures.....	57
13.3. Part 2 - Supplementary Tables and Figures.....	58
13.4. Part 3 - Supplementary Tables and Figures.....	58

1. Resumo da tese (Abstract in portuguese)

Os microARNs (miARNs) são pequenas moléculas de ARN com um tamanho de 20 a 24 nucleótidos (nt). Os genes dos miARNs são transcritos pela ARN polimerase II e os transcritos adquirem uma estrutura secundária em cadeia dupla em forma de gancho denominada de miARN primário (pri-miARN). As proteínas Dicer-like 1 (DCL1), HYPOSTATIC LEAVES (HYL1) e SERRATE (SE) formam um complexo enzimático que cliva o pri-miARN na região oposta à cabeça do gancho libertando uma das extremidades do miARN originando uma sequência em cadeia dupla mais curta denominada de miARN precursor (pre-miARN). As enzimas DCL1 e HYL1 voltam a interagir e a clivar a estrutura secundária no lado oposto do primeiro corte, libertando o miRNA em cadeia dupla (miARN maduro), sendo o seu complementar na cadeia oposta denominado de miARN*. A extremidade 3' de cada cadeia é metilada pela HUA ENHANCER 1 (HEN1), de forma a proteger o miRNA da degradação por uridilação. O miRNA maduro é então exportado do núcleo para o citoplasma pela proteína HASTY (HST). No citoplasma o miARN maduro associa-se à proteína Argonate 1 (AGO1), que integra o Complexo Silenciador Induzido por ARN, ocorrendo depois a dissociação e degradação do miARN*. O miARN direcciona o complexo enzimático para ARN mensageiros (mARN), que possuem uma região complementar, acabando por ter dois destinos, a clivagem ou repressão traducional. A clivagem no mARN, ocorre a meio da região complementar, entre os resíduos que estão emparelhados com os nucleótidos +10 e +11 do miARN guia (a contar da extremidade 5'). Os miARNs apresentam-se assim como reguladores negativos da expressão génica.

Em plantas a via de produção de miARNs e seu posterior funcionamento é regulada por um processo de "feed-back". As proteínas DCL1 e AGO1 são reguladas ao nível do mARN, pelo miR162 e pelo miR168 respectivamente. O principal objectivo deste trabalho foi analisar como se processa esta regulação em "feed-back" na planta modelo *Medicago truncatula* sujeita a condições de défice hídrico. Para tal foram utilizadas plantas adultas crescidas em vaso em condições controladas e posteriormente divididas em quatro grupos: plantas bem irrigadas (Controlo, Ct), plantas com supressão da irrigação por um período de 5 dias (Défice Hídrico Moderado, DHM), supressão da irrigação por um período de 8 dias (Défice Hídrico Severo, DHS) e plantas que após 8 dias de falta de irrigação foram regadas durante 3 dias (Recuperação, Rec). O estado hídrico das plantas foi avaliado por parâmetros fisiológicos: calculou-se o conteúdo hídrico relativo (CHR) das plantas de cada condição antes da sua recolha para armazenamento e posterior análise. Para cada condição, as plantas apresentaram em média valores de 80% para Ct, 50% para DHM, 30% para DHS e 80% para Rec. Com isto surgiu-nos a questão, será que ao nível celular/molecular as respostas estarão a ocorrer em paralelo às respostas fisiológicas? Para responder a esta questão escolhemos 4 genes descritos por serem induzidos em resposta ao défice hídrico em plantas, para comparar os seus valores de expressão com os de CHR. Os genes escolhidos foram: Early Responsive to Dehydration 1 (ERD1), Plasma membrane intrinsic protein 1;5 (PIP1;5), Dehydrin e L-Myo-inositol-1-phosphate Synthase (MI-1-PS). Todos aumentaram a sua expressão na parte aérea e nas raízes em condições de défice hídrico, apresentando uma relação inversa com o CHR da planta, à excepção do MI-1-PS, que na parte aérea reduziu a sua expressão em DHS em relação ao DHM. Desta forma, com os

parâmetros avaliados, conseguimos relacionar as respostas fisiológicas de *M. truncatula* com as respostas ao nível molecular e reforçar a confirmação do estado hídrico da planta avaliado pelo CHR. Para investigar o que ocorre na via de regulação em “*feed-back*” da produção/função dos miARNs em déficit hídrico, procedemos à análise da expressão de miR162 e miR168 em *M. truncatula* nos 4 estados hídricos. A sequência do miR162 para *M. truncatula* estava disponível na miRBase mas não a do miR168, o que nos levou a fazer a pesquisa deste miARN no genoma da planta. A validação de um miARN tem de obedecer a um mínimo de dois critérios: critério da biogénese e critério da expressão. Segundo o critério da biogénese, tem de ser identificado o precursor de um miRNA que consiga formar uma estrutura secundária em gancho estável. Para isto são avaliados vários parâmetros da estrutura secundária como a energia mínima necessária para adquirir aquela conformação, o conteúdo em adeninas e uracilos da sequência e o número de nucleótidos não complementares entre o miARN e seu miARN*. Após esta primeira análise no genoma de *M. truncatula*, foram encontrados dois possíveis genes MIR168 (denominados MIR168a e MIR168b), que dão origem a pre-miARNs com o mesmo miR168 (sequência de miR168a igual à de miR168b). Para verificar o critério da expressão, foram construídas sondas para miR162 e miR168a/b e procedeu-se à sua hibridação com ARN total (*northern blot*) isolado de diferentes tecidos de *M. truncatula*. Detectou-se expressão de ambos miRNAs, confirmando-se assim o critério de expressão para miR162 e miR168, que em conjunto com a confirmação do critério da biogénese nos permitiram avançar para a análise da sua expressão nas plantas em déficit hídrico. Foi usado o *northern blot* para avaliar a expressão do miR162 e miR168a/b em cada um dos estados hídricos tanto em raízes como em parte aérea. Observou-se que ambos aumentaram ligeiramente a sua expressão na parte aérea com o déficit hídrico. Na raiz o miR168 reduziu a sua expressão com o déficit hídrico mas no déficit hídrico severo a sua quantidade é maior que no moderado, continuando a ser inferior ao controlo. No caso do miR162, também se observa uma redução de expressão, mas superior à do miR168a/b, atingindo o nível mais baixo no DHS. A quantificação relativa (por *quantitative real time PCR*, qRT-PCR) do nível de expressão dos alvos destes miARNs (DCL1 e AGO1) nos diferentes tecidos e estados hídricos das plantas revelou um aumento da expressão de ambos na parte aérea e raízes em déficit hídrico. Este aumento foi maior para DCL1 nas raízes, estando os seus níveis negativamente relacionados com os níveis de expressão de miR162. Para o caso de miR168a/b e AGO1 não conseguimos estabelecer a mesma relação. Desta forma concluímos que DCL1 está a ser regulado nas raízes por miR162 em *M. truncatula* sujeita a condições de déficit hídrico.

Por último, fizemos a pesquisa bioinformática de outros genes codificantes de AGOs e DCLs no genoma de *M. truncatula*, usando os algoritmos de BLASTn e tBLASTn, com as sequências de mRNA e proteínas das DCLs e AGOs de *A. thaliana* como referência. Os domínios proteicos das sequências identificadas para *M. truncatula* foram analisados de forma a pesquisar a existência dos motivos característicos de cada família (DCLs ou AGOs). Foram identificados 3 genes codificantes para DCLs e 9 para AGOs. A análise de expressão por qRT-PCR foi feita para alguns destes genes nos tecidos e estados hídricos anteriormente referidos. Observámos que DCLs (MtDCL3) e AGOs (MtAGO4b, MtAGO4c e MtAGO7) envolvidas em vias de processamento de pequenos ARNs para além da dos miARNs têm os níveis de expressão aumentados em situação de déficit hídrico. A

MtAGO7 está envolvida na produção dos denominados *trans-acting small RNAs*, que estão associados ao desenvolvimento da planta ao regularerem a expressão de Auxin Response Factor genes (ARF) e a MtDCL3, MtAGO4b e MtAGO4c, estão envolvidas em processos de metilação do DNA e consequente silenciamento transcricional de genes.

2. Resumo (Abstract in portuguese)

Os microARNs (miARNs) são pequenas moléculas de ARN com um tamanho de 20 a 24 nucleótidos (nt) e apresentam-se como reguladores negativos da expressão génica. As proteínas Dicer-like 1 (DCL1) e Argonaute 1 (AGO1) estão envolvidas na via de produção de miARNs e seus mARNs são regulados por miR162 e miR168 respectivamente. O principal objectivo deste trabalho foi analisar como se processa esta regulação em “*feed-back*” negativo na planta modelo *Medicago truncatula* sujeita a condições de défice hídrico. As plantas foram sujeitas a progressiva supressão da irrigação e de acordo com a avaliação do conteúdo hídrico relativo (CHR) das folhas, foram divididas em quatro grupos: controlo, défice hídrico moderado, défice hídrico severo e recuperação. Primeiro, foi analisada a expressão de 4 genes descritos por serem induzidos em resposta ao défice hídrico - Early Responsive to Dehydration 1 (ERD1), Plasma membrane intrinsic protein 1;5 (PIP1;5), Dehydrin e L-Myo-inositol-1-phosphate Synthase (MI-1-PS) - para comparar com os valores do CHR das plantas. Todos aumentaram a sua expressão na parte aérea e nas raízes em condições de défice hídrico, apresentando uma relação temporal e inversa com o CHR das plantas, à excepção do MI-1-PS na parte aérea. Em segundo, a expressão de miR162 reduziu e de DCL1 aumentou com o défice hídrico. Desta forma concluímos que DCL1 está a ser regulado nas raízes por miR162 em *M. truncatula* sujeita a condições de falta de água. Em último, foram identificados em *M. Truncatula*, 3 genes codificantes para DCLs e 9 para AGOs envolvidos outras vias de processamento de pequenos ARNs. Observámos que têm os níveis de expressão aumentados em situação de défice hídrico. A MtAGO7 está envolvida na produção dos denominados *trans-acting small RNAs* e a MtDCL3, MtAGO4b e MtAGO4c, estão envolvidas em processos de metilação do DNA e consequente silenciamento transcripcional de genes.

Palavras chave: *M. truncatula*, Défice Hídrico, miR162, Dicer-like, Argonaute

3. Abstract

The microRNAs (miRNAs) are small RNAs with 20-24 nucleotides that negatively regulate the transcripts of several genes. Dicer-like 1 (DCL1) and Argonaute 1 (AGO1) are two key proteins involved in the miRNA biogenesis and their mRNAs are feedback regulated by miR162 and miR168 respectively. The main goal of this work was to study this mechanism of regulation of miRNAs biogenesis in *Medicago truncatula* under different water deficit conditions. Plants were subjected to progressive water deprivation and according to the evaluation of their water status by the leaf relative water content (RWC) were divided in four groups: control, moderate water deficit, severe water deficit and recovery. First, we analyzed the expression of 4 water deficit-responsive genes - Early Responsive to Dehydration 1 (ERD1), Plasma membrane intrinsic protein 1;5 (PIP1;5), L-Myo-inositol-1-phosphate Synthase (MI-1-PS) and Dehydrin - to establish a relation with the RWC of the plants. We found that the expression levels of ERD1, PIP1;5 and Dehydrin genes in plant shoots and roots and MI-1-PS in roots were temporally and inversely correlated with the plants RWC parameters, confirming that progressive water deficit induces an increase in transcription of water deficit-responsive genes. Second, we showed that miR162 was down-regulated in *M. truncatula* in water deficit whereas its target, DCL1, was up-regulated. This indicates that miR162 is regulating the mRNA levels of DCL1 in *M. truncatula* plants under water deficit. Third, in a bioinformatic screen we found and characterized the protein sequence of 3 and 9 new DCLs and AGOs in *M. truncatula*. The expression analysis of some of these genes under water deficit, showed up-regulation of genes encoding proteins involved in small RNA silencing pathways like direction of chromatin methylation for silencing of gene transcription (DCL3, AGO4b and AGO4c) or the production of trans-acting small interfering RNAs derived from TAS3 gene transcripts (AGO7).

Key words: *M. truncatula*, Water Deficit, miR162, Dicer-like, Argonaute

4. Introduction

4.1. *Medicago truncatula* as a model legume

Arabidopsis thaliana (L.) Heynh is considered a model plant but is a non-nitrogen-fixing plant, creating a flaw in the concept of model plant, making it important to develop a leguminous model with its ability to establish symbiotic interactions with rhizobia and mycorrhizae (Rose, 2008). With more than 18 000 species, legumes are the third largest family of higher plants (Young et al, 2003) and their ability to utilize atmospheric nitrogen fixation by the rhizobia symbiosis is of key importance in the biosphere and in agricultural systems (Rose, 2008).

In 1990, Barker et al (1990) proposed *M. truncatula* as a model legume for the study of the molecular genetics of the rhizobia-legume symbiosis. *M. truncatula* (2n=16) is diploid and autogamous, has a relatively small genome (~500 mbp, Bennett and Leitch, 1995) and a generation time of ~3 months in long day conditions (Barker et al, 1990).

Several tools for molecular genetics and genomics have been recently developed. Genome sequencing began in 2002 (Young et al, 2005) and has continued until the present time (<http://www.medicago.org/genome/>, accessed 18 September 2009). Expressed sequence tags (ESTs) (259642 ESTs) were sequenced and have been assembled into 29673 TCs (tentative consensus sequences) and 26696 singleton ESTs (<http://compbio.dfci.harvard.edu/tgi/>, accessed 18 September 2009). *M. truncatula* is able to be transformed by *Agrobacterium tumefaciens* and to regenerate via somatic embryogenesis (Araujo et al, 2004).

4.2. Plant Responses to Water Deficit

Plant responses to water deficit are multiple and interconnected, ranging from physiological, biochemical/cellular and molecular responses.

4.2.1. Physiological and biochemical responses

One of the first responses to water deficit is the closure of stomata to prevent the transpiration water loss (Hsiau et al, 1973; Taiz and Zeiger, 2002). Nunes et al (2008) showed that under severe water deficit conditions, *M. truncatula* resistance to water shortage involved mainly water deficit avoidance mechanisms through a decrease in stomatal conductance, caused by stomatal closure. Stomatal closure in response to a severe water deficit stress primarily results in a decline in the photosynthesis rate (Taiz and Zeiger, 2002; Mahajan et al, 2005). A very severe water deficit condition results in limited photosynthesis due to decline in Rubisco activity (Bota et al, 2004).

A water deficit consequence in plants is the shoot growth inhibition (Chaves et al, 2003) and cell division and expansion are inhibited by long exposure to mild water stress (Hsiau et al, 1973). Nutrient uptake by plants is decreased under water deficit stress conditions due to reduced transpiration, impaired active transport and membrane permeability resulting in reduced root absorbing capacity (Tanguilig et al, 1987).

A biochemical/cellular response of the plants to water deficit is known as osmotic adjustment, and consists in the decrease of the cellular osmotic potential by the accumulation of solutes (Mahajan et

al, 2005). The osmotic compounds synthesized include proline, glycine betaine and trehalose (Taiz and Zeiger, 2002; Seki et al, 2007).

4.2.2. ABA signaling

Water deficit triggers the production of the phytohormone abscisic acid (ABA), which in turn causes stomatal closure and induces expression of stress-related genes (Shinozaki and Yamaguchi-Shinozaki, 2007). Several water deficit-inducible genes are induced by exogenous ABA treatment (ABA-dependent), whereas others are not affected (ABA-independent). Gene transcription is regulated through the interaction of transcription factors with specific regulatory sequences in the promoters of the genes they regulate (Taiz and Zeiger, 2002). Two of the cis-acting elements present in promoters of water deficit regulated genes are ABRE (ABA-responsive element), and DRE (dehydration-responsive element)/CRT (C-Repeat), both of which are involved in stress-inducible gene expression. ABRE with the DNA-binding motif ACGTGGC, is the major regulatory region in ABA-responsive gene expression. For ABA-responsive transcription, a single copy of ABRE is not sufficient (Yamaguchi-Shinozaki and Shinozaki, 2006). Two ABRE motifs or one ABRE and its coupling element are required in the promoter region for the expression of ABA-responsive genes. The basic leucine zipper (bZIP) transcription factors, AREB/ABF, can bind to ABRE, thereby activating ABA dependent gene expression (Choi et al, 2000; Uno et al, 2000).

Transcription factors belonging to the ERF/AP2 family that binds to these DRE/CRT elements were isolated and termed CBF/DREB1 and CBF/DREB2. Their conserved DNA-binding motif is A/GCCGAC (Shinozaki and Yamaguchi-Shinozaki, 2007).

4.2.3. Molecular Responses

Proteins denominated as aquaporins, are important in water flux control in cells (Shao et al, 2008). Several protease genes are also induced and they degrade other proteins that are denatured by stress (Bray, 1993; Ingram and Bartels, 1996). The expression levels of ubiquitin genes are up-regulated under water deficit, and ubiquitin enzyme tags proteins that are targeted for proteolytic degradation. Chaperones involved in protein repair by helping other proteins to recover their native conformation after denaturation or misfolding during water stress, are induced too. Other group of genes regulated by water deficit is the LEA (late embryogenesis abundant) proteins (Taiz and Zeiger, 2002). These proteins are suspected to have a role in the cellular membrane protection and in preventing the crystallization of cellular proteins and other molecules.

4.3. What are miRNAs?

Although most genes use RNA in the form of messenger RNA (mRNA) as a coding intermediate for protein production, there are many genes whose final products are RNA (Bartel and Bartel, 2003). These noncoding RNAs range from the transfer and ribosomal RNAs to the 18–25 nucleotide (nt)-long regulatory small RNAs (sRNAs) that can mediate endogenous gene silencing at both the transcriptional and pos-transcriptional levels (Bartel and Bartel, 2003; Voinnet et al, 2009). sRNAs are classed into two categories based on their biogenesis: small interfering RNAs (siRNAs) are processed

from long, perfectly double-stranded RNA, and microRNAs (miRNAs) from single-stranded RNA transcripts (transcribed from MIR genes) that have the ability to fold back onto themselves to produce imperfectly double-stranded stem loop precursor structures (Eamens et al, 2008).

MicroRNAs (miRNAs) are 20-24nt RNA molecules and fundamental sequence-specific post-transcriptional regulatory elements of eukaryotic genomes (Voinnet et al, 2009). The first miRNA, lin-4, was discovered in *Caenorhabditis elegans* in 1993 (Lee et al, 1993), and today hundreds of miRNAs have been identified in plants and animals, including several hundred unique miRNAs in Arabidopsis alone (Millar and Waterhouse, 2005).

4.4. miRNA Biogenesis

4.4.1. miRNA Transcription

Plant MIR genes are located generally in intergenic regions and are transcribed into a primary precursor known as primary miRNA (pri-miRNA) (Figure 1) (Jones-Rhoades and Bartel, 2004, Voinnet, 2009). MIR genes are class-II genes, that is, genes transcribed by RNA polymerase II (pol II) (Lee et al, 2004, Xie et al, 2005). For instance, primary transcripts of some microRNA genes contain poly(A) tails, or the cap structure (Tam, 2001; Aukerman and Sakai, 2003). The great majority of the pri-miRNAs begin with an adenosine, which is located within 40nt downstream of a conserved TATA box-like sequence (Xie et al, 2005). A bioinformatic analysis of 800nt regions upstream of the mapped transcription start sites in these genes identified binding motifs for a number of known transcription factors (Megraw et al, 2006). Expressions of some microRNA genes are regulated by phytohormones, like ABA, indicating that some MIR genes possibly have ABRE regulatory motifs (Liu et al, 2009a).

4.4.2. Processing to mature form

The pri-miRNA transcript is cleaved by DCL1 in the nucleus, to produce the shorter precursor miRNA (pre-miRNA) dsRNA molecule (Kurihara and Watanabe, 2004; Kurihara et al, 2006). The first DCL1-catalyzed cleavage step in the miRNA biogenesis pathway is made just below the miRNA duplex region of the dsRNA stem loop (Figure 1) (Voinnet, 2009). pri-to-pre-miRNA conversion also necessitates the double-stranded RNA-binding protein HYPONASTIC LEAVES1 (HYL1) and the C2H2-zinc finger protein SERRATE (SE), which interact with DCL1 in nuclear processing centers called D-bodies (Fang and Spector, 2007; Kurihara et al, 2006). The miRNA duplex is then released from the pre-miRNA stem loop structure by the second cleavage step of the miRNA pathway, which is again directed by the combined action of DCL1 and HYL1 (Wu et al, 2007; Liu et al, 2005).

4.4.3. Methylation

The two-nucleotide 3' overhangs of the liberated miRNA duplex are methylated by the sRNA-specific methyltransferase HUA ENHANCER1 (HEN1) (Figure 1) (Chen et al, 2002; Yu et al, 2005). The duplexes of siRNAs are also methylated by HEN1, a process that appears to be plant specific and are assumed to protect all sRNA species from polyuridylation and degradation (Li et al, 2005; Yang et al, 2006). The recent isolation of Arabidopsis exonucleases that degrade miRNA single strands in vitro further highlights the importance of miRNA stability control in plants (Ramachandran and Chen, 2008).

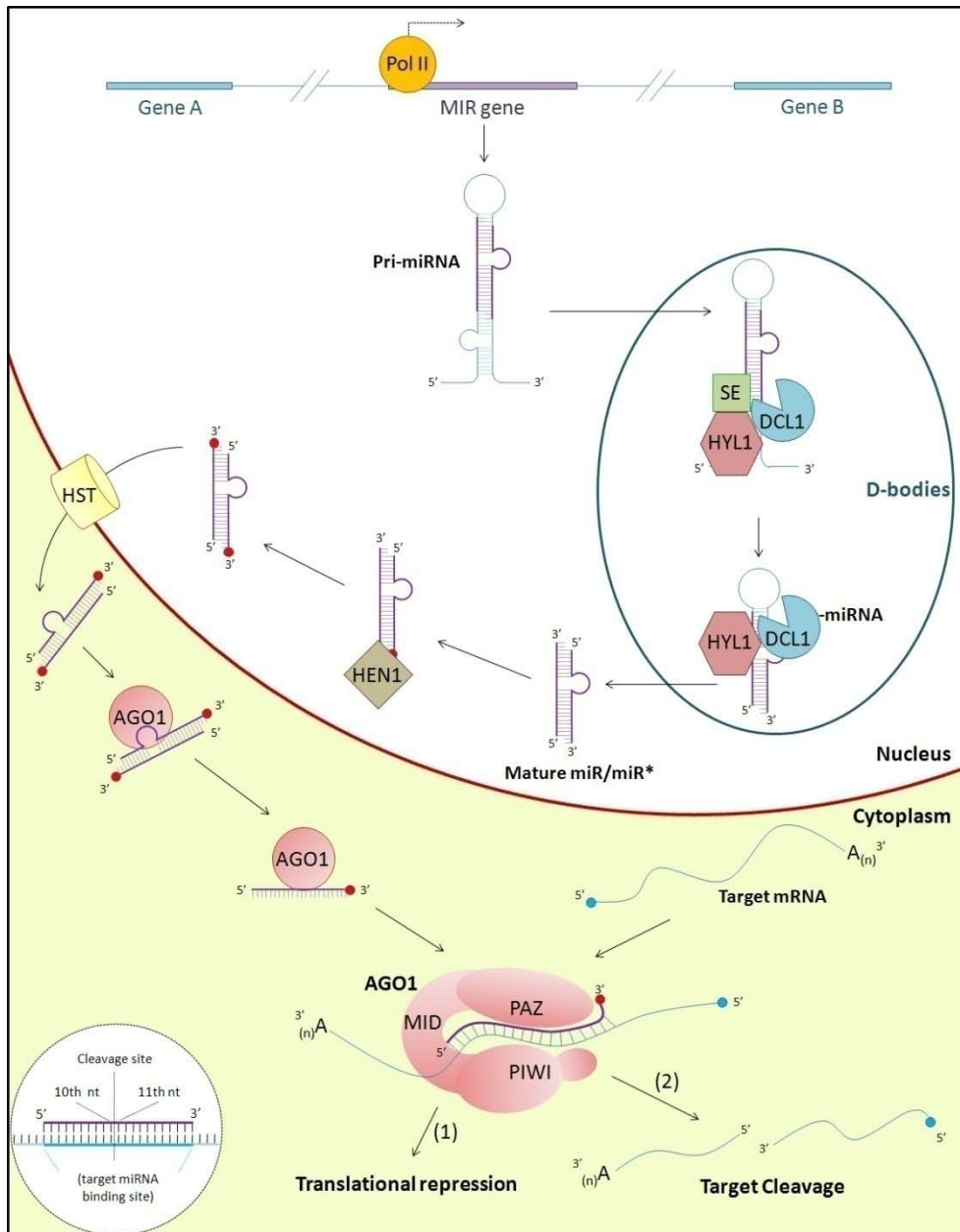


Figure 1: miRNA maturation and target cleavage in Plants - a simplified schematic. Plant MIR genes are transcribed by RNA Polymerase II (Pol II). Transcripts fold into a hairpin structure denominated as primary miRNA (pri-miRNA). The Dicer-like 1 (DCL1), HYPOSTATIC LEAVES (HYL1) and SERRATE (SE) proteins form a complex DCL1 that cleaves it into a shorter precursor miRNA (pre-miRNA) dsRNA molecule in the D-bodies. The DCL1 and HYL1 are involved in the second cleavage of the pre-miRNA into the mature miRNA. The 3' ends of miRNAs and the complementary strands (miRNA*) are methylated by HUA ENHANCER 1 (HEN1), which protects against uridylation and subsequent decay. The nuclear exporter HASTY (HST) exports the mature miRNAs to the cytoplasm where one strand of miRNA duplexes associates with AGO proteins. Mature miRNAs bind to Argonaute1 (AGO1) protein integrated in the RNA-induced silencing complex (RISC) and directs the cleavage of partially complementary target mRNAs (1) or target translation repression (2). AGO proteins have three domains: MID, PAZ, and PIWI. It has been proposed that the MID domain binds miRNA 5' ends while the PAZ domain binds miRNA 3' ends to position miRNAs for PIWI-mediated cleavage, of partially complementary mRNA target sites. The cleavage site is between the nucleotides 10 and 11 starting from the 5' end of the miRNA sequence.

4.4.4. Export from Nucleus to Cytoplasm

In the plant cells the miRNA duplex is then transported to the cytoplasm from the nucleus, with several classes of miRNAs relying on the action of the HASTY (HST) protein one ortholog of the animal Exportin-5 (Figure 1) (Park et al, 2005). However, the exact role of HST in miRNA biogenesis remains unclear, as other families of miRNA appear to be transported to the cytoplasm via a HST-independent mechanism (Park et al, 2005).

4.4.5. RISC assembly

The methylated miRNA/miRNA* duplex strand is then incorporated into a protein complex called RNA-induced silencing complex (RISC) whose major protein component is an Argonaute1 (AGO1) protein (Voinnet, 2009). In humans and Drosophila, the complementary strand of the miRNA (miRNA*) is cleaved by AGO2 and the cleaved fragments are released, resulting in the formation of RISC with one miRNA strand (Matranga et al, 2005; Miyoshi et al, 2005).

The determination of which strand of the duplex ends up in RISC is largely based on the thermodynamic properties of the two ends of the miRNA/miRNA* duplex. The strand in which the 5' end is less thermodynamically stable in the duplex becomes preferentially incorporated into RISC (Khvorova et al, 2003; Schwarz et al, 2003). Plant miRNAs and siRNAs appear to follow this asymmetry rule in RISC loading. In fact, most plant miRNAs begin with a U residue such that the 5' end of the miRNAs tends to be engaged in A–U, rather than G–C, hydrogen bonding (Reinhart et al, 2002; Rajagopalan et al, 2006).

4.5. Mode of action of Plant miRNAs

Plant miRNAs have a high degree of sequence complementarity to their target mRNAs and direct the cleavage of the target mRNAs in the middle of the complementary regions, between nucleotide +10 and +11 starting from the 5' end of the miRNA sequence (Figure 1) (Llave et al, 2002; Tang et al, 2003). Cleaved RNA fragments mapping to central regions of predicted hybrids were indeed retrieved in experiments using the technique of rapid amplification of 5' complementary DNA ends (5' RACE).

Brodersen et al (2008) showed that the action of most plant miRNAs commonly involves a combination of target degradation and translational repression, which affects the pool of transcripts remaining after cleavage. Importantly, neither the position (ORF, 5' or 3'UTR) nor the degree of pairing at the miRNA target sites appeared to be predictive of the prevalence of one process over the other. Furthermore, 5' RACE, the technique predominantly used to validate miRNA targets, is a qualitative procedure and does not indicate the extent of target cleavage (Voinnet, 2009).

4.6. miRNAs discovery in Plants

Three major methods have been used for miRNA discovery: forward genetics, direct cloning and sequencing, and bioinformatic prediction (Chen, 2008).

Forward genetic studies, seeks to find the genetic basis of a phenotype. Despite having resulted in the identification of a few miRNAs, this method provides hints to the, functions of these miRNAs in addition to their isolation (Chen, 2008).

A second approach to miRNA discovery is bioinformatic prediction that employs several different algorithms to predict miRNAs. The features that the algorithms search for in the genomic sequences are based on our current knowledge of plant miRNAs (Bonnet et al, 2004; Jones-Rhoades and Bartel, 2004; Adai et al, 2005). These features include the intergenic location of the MIR genes, the high degree of sequence complementarity of miRNAs to their mRNA targets, the hairpin structures of the precursors, and the conservation of some miRNAs between species.

A third method for miRNA discovery is direct cloning and sequencing. One advantage of this method is that the antisense strands of the small RNAs (miRNA* or siRNA*) that are also released from the precursor during small RNA biogenesis are often detected, although at much lower frequency compared to the sense strands. This also helps to distinguish the miRNAs from the small RNA populations since, by definition, the miRNA and the miRNA* should map to the same strand and should be separated by hundreds of nucleotides (Chen, 2008).

4.7. miRNA function in Plants

4.7.1-Regulatory functions of miRNAs in development

At least three types of regulatory functions are specified for miRNAs in plant development:

4.7.1.1. Spatial expression

Some miRNAs define the spatial expression pattern of their targets, that is, miRNAs control developmental processes by depleting the tissue where they accumulate from their target, thereby regulating cell fate transition or differentiation. This function is usually proposed when the target expression pattern increases in the absence of miRNA regulation (Garcia, 2008). Spatial restriction was proposed for the action of miR165/166 in the regulation of PHB in Arabidopsis and RLD1 in maize. Those genes are polarly expressed on the adaxial side of the leaf. In contrast, *in situ* hybridization in maize and Arabidopsis revealed the accumulation of miR165/166 on the adjacent abaxial side (Kidner and Martienssen, 2004; Juarez et al, 2004).

4.7.1.2. Buffering Function

Other miRNAs function to prevent excessive variations in the accumulation level of a target on a transcriptionally defined expression domain. This action prevents fluctuations and accurately controls the expression level of targeted genes (Garcia, 2008). miR156 exerts a quantitative control on the accumulation of its target SPL9 that determines the timing of leaf initiation. The expression domains of miR156 and SPL9 are largely overlapping and the expression of SPL9 in plants with reduced miR156 activity increases, and do not accumulate ectopically (Wang et al, 2008).

4.7.1.3. Temporal regulation

This kind of regulation involves opposite gradients of miRNA and target genes accumulation up to a predetermined target threshold inducing developmental transition (Garcia, 2008). miR172 targets AP2-like genes and controls the plant transition from vegetative to reproductive phase (Aukerman and Sakai, 2003). The progressive accumulation of miR172 during vegetative development induces a

gradual decrease in the abundance of its targets. When the target reaches a critical threshold, it promotes flowering (Chen, 2004).

4.7.2. miRNAs involved in plant response to biotic and abiotic stress

The involvement of miRNAs in plant response to pathogens was shown in the case of miR393 that plays a crucial role in the plant defense against pathogens by targeting Transport Inhibitor Response 1 (TIR1), an auxin receptor (Navarro et al, 2006). Several miRNAs have an important role in plant response to abiotic stress like: mechanical stress (Lu et al, 2005), oxidative stress (Sunkar et al, 2006), phosphate starvation (Fujii et al, 2005), UV-B stress (Zhou et al, 2007), salinity and cold stress and water deficit (Sunkar et al, 2004; Zhao et al, 2007).

4.8. Feedback Regulation

miRNA biogenesis is under feedback regulation such that two key players in miRNA biogenesis and function are themselves regulated by miRNAs (Figure 2). DCL1 mRNA has a binding site for miR162, which leads to the cleavage of DCL1 mRNA (Xie et al, 2003). miR168 has a binding site in AGO1 mRNA and leads to AGO1-mediated cleavage of AGO1 mRNA (Vaucheret et al, 2004).

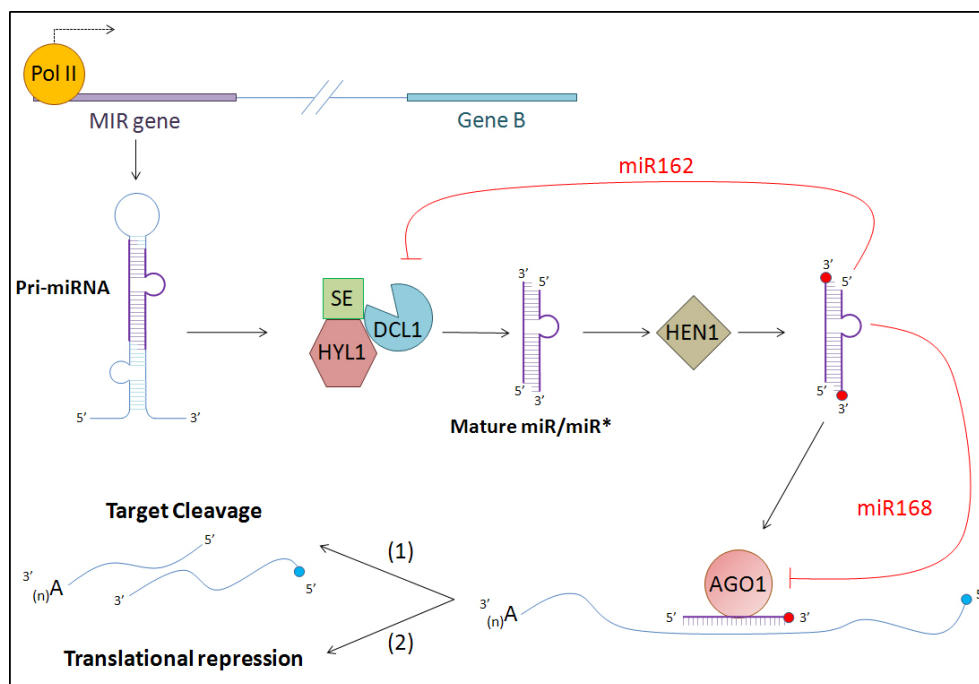


Figure 2: miRNA biogenesis feedback regulation and major steps in miRNAs biogenesis. Primary miRNA transcripts (pri-miRNAs) are processed to miRNA/miRNA* duplexes by the coordinated activities of DCL1, HYL1 and SE within the nucleus. Both strands of the duplex are methylated on the 3' end by HEN1. In the cytoplasm the miRNA strand becomes incorporated into an AGO1-containing RNA-induced silencing complex (RISC) and directs either mRNA cleavage (1) or translational inhibition (2). miRNA feedback regulatory loops are depicted in red: miR162 and miR168 target DCL1 and AGO1 mRNAs, respectively.

5. Objectives

To perform the study in water deficit conditions our plant material consisted in 4 groups of *M. truncatula* plants, each one subjected to a different water deficit treatment. The plants water status was evaluated by a physiological parameter named relative water content (RWC).

The **first** objective of this work was to find a relation between the RWC of the plants and the molecular responses of the plants cells. For this, we quantified the expression levels of 4 genes previously reported to change their expression levels in water deficit and established a relation with the RWC values.

Our **second** objective in this work consisted the studying of the feedback regulation of the biogenesis of the miRNAs under water deficit conditions in *M. truncatula*. The elements targeted for expression analysis were the miR162 and miR168 that regulate the two key enzymes of the miRNA biogenesis, DCL1 and AGO1 respectively.

The **third** objective was to search for other Dicer-like and Argonaute genes in the *M. truncatula* genome and to analyze their expression in the 4 water deficit conditions, to know if they could be involved in plant response to water deficit.

6. Methods

6.1. Plant material and treatment conditions

6.1.1. Plant material and growth conditions

Medicago truncatula Gaertn. cv. Jemalong seeds were scarified in concentrated anhydrous sulphuric acid and sterilized according to Araújo et al (2004). Seeds were placed in Petri dishes with soaked filter paper at 24°C in dark conditions for 3 days and then transferred to the growth chamber (thermoperiod of 25/18°C, day/night, relative humidity of 40%) to germinate within 2–3 days still in dark conditions. Seedlings were allowed to grow for one week under light conditions (photoperiod of 16/8 h, day/night, and a Photosynthetic Photon Flux Density (PPFD) of around 500 μmol m⁻² s⁻¹ obtained from white fluorescent and incandescent lamps, TLD58W/77, Halo Star 64476, L30W/77, L58W/77. Seedlings were transferred to trays with vermiculite for 2 weeks and then individually transferred to 0.5L pots with standard commercial soil (“terra de Montemor”, Horto do Campo Grande, Lisboa, Portugal). The plants were kept in the growth chamber, under the same conditions, through the entire experiment, even during water deficit imposition.

6.1.2. Water deficit treatments

Eight-week-old plants were divided in 4 groups, corresponding to the different levels of soil water availability imposed to the plants. The Control group (Ct) was constituted by plants maintained fully irrigated with tap water until the maximum soil water capacity. For the other 3 groups, the Soil Water content (SWC) was decreased by suppression of soil irrigation for approximately 5 days (Moderate Water Deficit, MWD) and 8 days (Severe Water Deficit, SWD) and rewatering took place for 3 days, after 8 days of water suppression, until plants were recovered (Rec). Each group was constituted by plants taken at day 5 and 8 after water suppression and day 3 after rewatering. Leaf water status was monitored by determination of the plant Relative Water Content (RWC), as described in Catsky (1960). Briefly, the weight of one leaf was determined immediately after cutting (Fresh Weight, FW). Each leaf was then submerged in water, at room temperature for 24h in the dark (Turgid Weight, TW). Finally, the leaves were dried in a 70°C chamber for at least 48h or until they were completely dried (Dried Weight, DW). RWC was calculated according to the formula: $RWC = (FW - DW) / (TW - DW) \times 100$. The youngest fully expanded leaf of the main branch was used for all measurements. Shoot and roots of water deficit treated and control plants were stored at -80°C after evaluation of plant relative water content. Average RWC in Ct plants was around 80%. Water deficit treated plants had RWC of about 50% (MWS) and 30% (SWS). Recovered plants showed RWC of about 30% after 8 days of water deprivation and then recovered to RWC of 80% after 3 days of watering.

6.2. RNA extraction

6.2.1. RNA Extraction from *Medicago truncatula*

1. Plant tissues (about 100mg) were frozen in liquid nitrogen and ground into fine powder using mortar and pestle;
2. 700 μl of extraction buffer (8.0M guanidine hydrochloride, 20mM 2-(N-morpholino)ethanesulfonic acid (MES), 20mM ethylenediaminetetraacetic acid (EDTA), 2% β-

mercaptoethanol, pH7.0)(van Kan et al, 1992) were added and after vigorous vortexing for 30s an equal volume of phenol (pH5.2) was added, vortexed vigorously and centrifuged at 12000rpm for 10min at room temperature;

3. The supernatant was transferred to a fresh tube containing one equal volume of phenol:chloroform:isoamyl alcohol (25:24:1) vigorously vortexed and centrifuged at 12000rpm for 2min;
4. The upper phase was removed to a fresh tube and step 3 was repeated again (to improve the quality of RNA extraction);
5. The upper phase was removed to a fresh tube containing equal volume of chloroform:isoamyl alcohol (24:1), vigorously mixed and centrifuged at 12000rpm for 2min;
6. The upper phase was transferred to a fresh tube containing 2 volumes of absolute ethanol and 1/20 volumes of 4M sodium acetate (pH5.2), mixed gently and left at -20°C for 30min.
7. After centrifugation for 10min at 12000rpm, the RNA was washed with ethanol 70%, air-dried on ice and resuspended in 100µl RNase-free water.

6.2.2. RNA integrity and quantification

The integrity of the RNA was verified running it in a 2% agarose gel (0,5X Tris-Borate-EDTA Buffer (TBE) (10X solution: 890mM Tris Base, 890mM Boric acid, 20mM EDTA)), stained with Sybr Safe (Invitrogen, Carlsbad, California, USA) (Supplementary Figure 1). The RNA was previously denatured in formamide buffer (10ml dionized formamide, 200µl 0,5M EDTA (pH 8.0), 10mg xylene cyanol, 10mg bromophenol blue) at 65°C for 10min. RNA quantification was done in a NanoDrop ND1000 (Thermo Scientific, Wilmington, Delaware, USA).

6.3. miRNA prediction

6.3.1. Bioinformatic prediction of precursor and mature miR168 gene

The precursor and mature sequences of *Glycine max* miR168 were downloaded from the miRBase (Figure 3) (<http://microrna.sanger.ac.uk/sequences/>, Griffiths-Jones et al, 2006). The algorithm BLASTn was used to search for similar sequences in the *M. truncatula* genome databases, at the National Center for Biotechnology Information (NCBI) (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>), at the *Medicago truncatula* Gene Index (MTGI) 9.0 (<http://compbio.dfc.harvard.edu/tgi/>) and at the *Medicago truncatula* Genome Sequencing Project (MTGSP) (*Medicago truncatula* genome release version 2.0 or Mt2.0) (<http://www.medicago.org/genome/>) using the CviT-Blast. To select the putative miR168 sequences for further analysis two criteria were considered: (1) good similarity with the precursor of *G. max* miR168 sequence; (2) no more than four mismatches with the *G. max* mature miR168 (Zhang et al, 2006a).

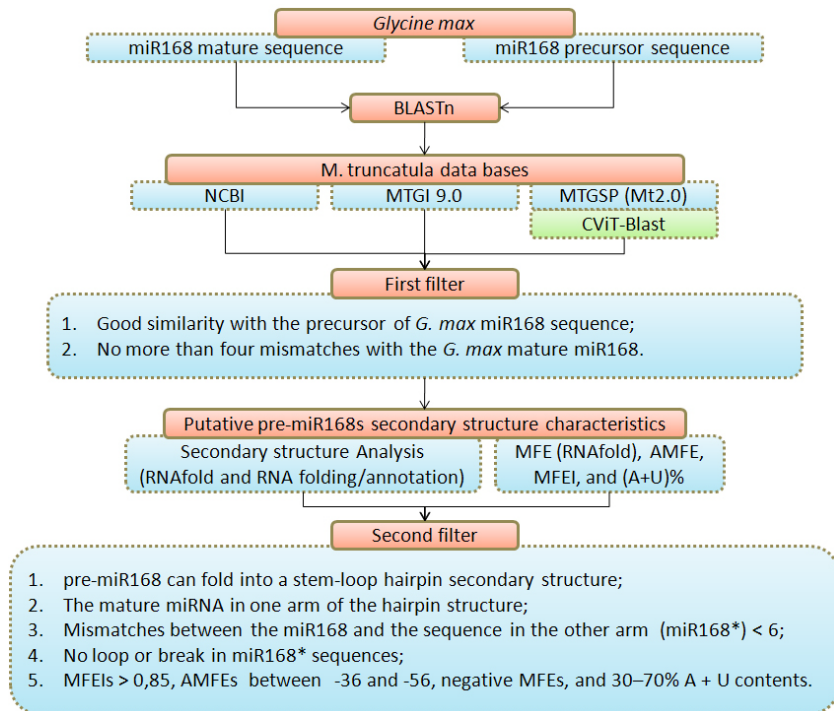


Figure 3: Simplified scheme of the bioinformatic search of MIR168 gene(s) in *M. truncatula*. The parameters used to evaluate the putative mature and precursor miR168 are shown (first and second filter). NCBI, National Center for Biotechnology Information; MTGI, Medicago truncatula Gene Index 9.0; MTGSP, Medicago truncatula Genome Sequencing Project (Medicago truncatula genome release version 2.0 or Mt2.0).

6.3.2. Characterization, secondary structure, and targets of miR168

The length, composition of AU and genomic location were determined for the putative pre-miR168 found for *M. truncatula*. The pre-miRNA *plus* 18bp to the 3' side and to the 5' side (Ambros et al, 2008) were used to obtain the minimum free energy (MFE) (the smallest energy necessary for the folding of the RNA molecule). The MFE and the most favorable RNA secondary structure for the sequences were determined via RNAfold software included in the Vienna RNA Package 1.8.2 (<http://rna.tbi.univie.ac.at/>) (Hofacker et al, 2003). For the secondary structure prediction was used too the "RNA folding/annotation" tool (<http://srna-tools.cmp.uea.ac.uk/>) (Moxon et al, 2008).

The adjusted MFE (AMFE) was calculated applying the equation (Zhang et al, 2006b):

$$AMFE = (MFE \div \text{Lenght of precursor used to calculate MFE}) \times 100$$

(Equation 1)

The minimum folding free energy index (MFEI) (Zhang et al, 2006b) for the sequences was determined using the formula:

$$MFEI = AMFE \div (G + C)\%$$

(Equation 2)

RNA sequences were considered miRNA candidates only if they fit the following criteria (Zhang et al, 2006a): (1) a RNA sequence can fold into an appropriate stem-loop hairpin secondary structure; (2) a mature miRNA sequence site in one arm of the hairpin structure; (3) miRNAs had less than six mismatches with the opposite miRNA* sequence in the other arm; (4) no loop or break in miRNA* sequences; (5) predicted secondary structures had MFEs higher than 0.85, AMFEs with ΔG between -36 and -56, negative MFEs, and 30–70% A + U contents (Zhang et al, 2006a, Zhang et al, 2006b). The targets of the miRNAs were determined using the “Target prediction tool” available online (<http://srna-tools.cmp.uea.ac.uk/>) (Moxon et al, 2008).

6.4. Identification of genes and their characterization

6.4.1. Identification of genes in *M. truncatula* by homology to *A. thaliana*

The mRNA and protein sequences of *Arabidopsis thaliana* Early Responsive to Dehydration 1 (ERD1), Plasma membrane intrinsic protein 1;5 (PIP1;5), Argonautes and Dicer-like genes were downloaded from the NCBI (<http://www.ncbi.nlm.nih.gov/>) (Supplementary Table 1). The algorithms BLASTn and tBLASTn were used to search the nucleotide sequence of the genes of interest at the NCBI database and the MTGSP (Mt2.0). The analysis performed in the MTGSP database were done in the genome sequenced to the data (version 2.0) using CVIT-Blast and in the annotated genome (version 2,0) using IMAG-Blast. In parallel, the same algorithms were used to search in the MTGI 9.0. The Argonautes and Dice-like genes results were analyzed in the following way (Figure 4): (1) only the mRNA sequences with more than 1000bp were accepted; (2) identification of the same sequences that were obtained in different databases (NCBI Vs MTGSP Vs MTGI); (3) the sequences that after BLASTn against *A. thaliana*, in NCBI, did not give a protein belonging to the Argonaute and Dicer-like families were discarded. Steps 1 and 2 were also used for the ERD1 and PIP1;5 identification.

6.4.2. Characterization of the *Medicago truncatula* Argonaute and Dicer-like genes

The protein and nucleotide sequences of several DCLs and AGOs genes were downloaded from MTGSP and MTGI. In the cases that the protein sequence was not available, the translation of the nucleotide sequence was done with the Translate software from Expert Protein Analysis System (ExpASY) (<http://www.expasy.ch/>). The end of translation was considered when the first stop codon appeared.

The newly identified genes in this study were named based on the nomenclature used in *A. thaliana* and on the basis of their phylogenetic relatedness to other members of the same family.

Determination of the genes Open Reading Frames (ORFs) sizes was done and characteristics of the proteins like size and isoelectric point (Pi) were done with the Protein Isoelectric Point software and calculation of molecular weight (MW) was done using the Protein Molecular Weight software, both softwares from the Sequence Manipulation Suite (SMS) package (version 2.0) (<http://www.bioinformatics.org/sms2/>).

The chromosomal location of the genes and their locus was determined doing BLASTn in MTGSP database.

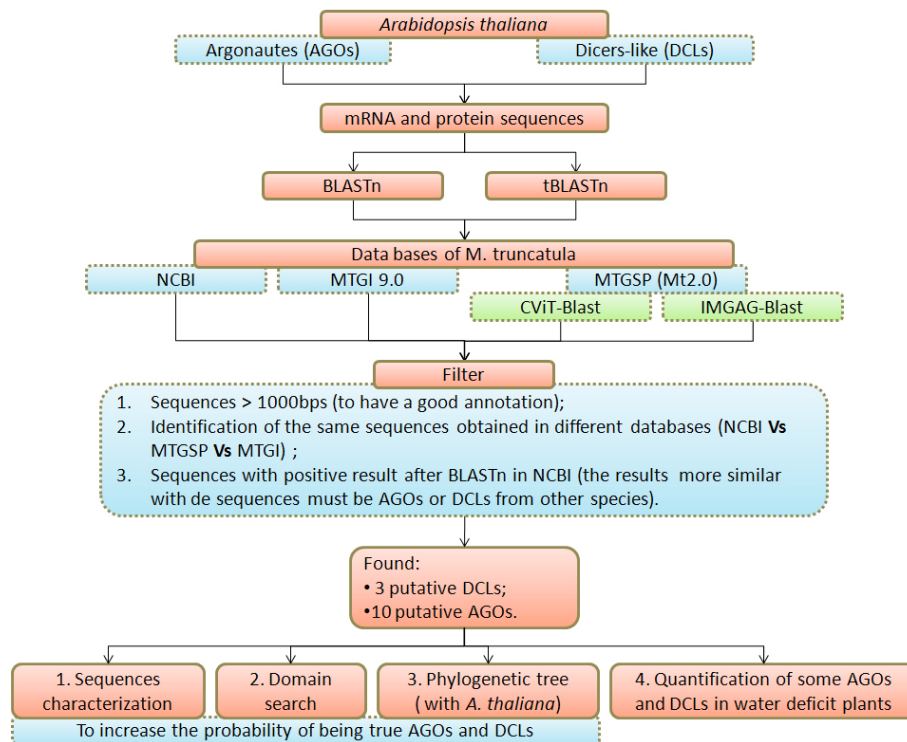


Figure 4: Scheme of the bioinformatic search for the Argonautes and Dicer-like genes in *M. truncatula* genome. The selection of the most provable sequences to be AGOs and DCLs was done performing a filter selection of the sequences. Was done a genomic, mRNA and protein sequence characterization, the protein domains were found using NCBI Conserved Domain Database (NCBI-CDD) and the phylogenetic relatedness between the proteins of *M. truncatula* and *A. thaliana* determined. BLASTn performed in the NCBI against *A. thaliana*. NCBI, National Center for Biotechnology Information; MTGI, *Medicago truncatula* Gene Index 9.0; MTGSP, *Medicago truncatula* Genome Sequencing Project (Mt2.0).

6.4.3. Protein domain search

The domain search was performed in NCBI Conserved Domain Database (NCBI-CDD). (Marchler et al, 2009) The characteristic catalytic domain of AGOs protein was obtained from the alignment of the PIWI domains from *M. truncatula* and the known amino acid positions of AGO1 from *A. thaliana*.

6.5. Alignment and phylogenetic trees

6.5.1. Nucleotide or amino acid alignment and phylogenetic tree build

The protein alignment was done using CLUSTALW software (Thompson et al, 1994) present in MEGA4.0 (Tamura et al, 2007). The phylogenetic trees based in protein sequences were generated using the following parameters: the distance model for amino acid substitution was Jones-Taylor-Thornton (JJT) matrix, the type of tree the Neighbor-Joining, and was performed a bootstrap analysis with 1000 replications. Pre-miRNA sequence alignment was performed using RCOFFEE software that do the alignment taking in account the secondary structure of the RNA molecule (<http://www.tcoffee.org/>) (Notredame et al, 2000). The phylogenetic tree was built using: the distance model for nucleotide substitution Maximum Composite Likelihood, the Neighbor-Joining tree and bootstrap analysis with 1000 replications. The software used for tree construction on both cases was MEGA4.0 (Tamura et al, 2007).

6.6. Quantitative Real Time PCR (qRT-PCR)

6.6.1. Quantitative PCR assay

The RNA samples were treated with the TURBO DNA-free Kit (Ambion, Austin, Texas, USA) to eliminate DNA contaminations (Figure 5). A pool of RNA (four plants) was made by treatment and organs (shoot and root). 1ug of RNA from each pool was reverse transcribed using the Promega - ImProm-II™ Reverse Transcription System (Promega, Madison, Wisconsin, USA) according to the manufacturer's instruction. Three independent reverse-transcription reactions (RT) were performed using the RNA pools and each one was diluted 5-fold (~10ng of initial RNA) before each quantitative Real Time Polymerase Chain Reaction (qRT-PCR) reaction.

PCR primers (Supplementary Table 2) were designed using the Beacon Designer software (version 7.0) (Premier Biosoft International, Palo Alto, California, USA). Reference genes were selected based on a pre-screen of HDA3, L2, APRT, ELF-1 α , ACT7 and ACT11 genes on different treatments and plant tissues (shoots and roots) using the geNorm (Vandesompele. et al, 2002) and NormFinder (Anderson et al, 2004) in Genex software (version 4.3.8) (MultiD, Göteborg, Sweden). Primers were designed to have an optimal size of 22bp (18–24bp), GC content of 40–60% and melting temperature (T_m) of 58–62°C. The AGO1 and DCL1 primer pair was designed to give one amplicon with the cleavage site of miR168 and miR162 respectively. This way, the product of amplification just derives of the cDNA from mRNA not cleaved by the respective miRNA. Other criteria, such as primer self-annealing, were also taken into account. Predicted fragment size ranged between 80 and 180bp. Oligonucleotides were synthesized by Stabvida (Stabvida, Oeiras, Portugal).

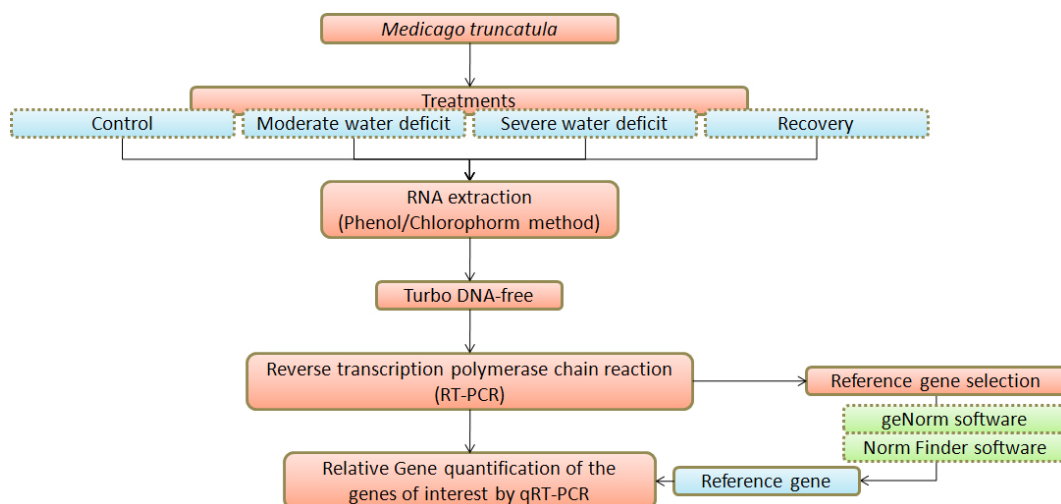


Figure 5: Synthetic scheme of the main steps starting from RNA extraction to quantification of gene expression on *M. truncatula*.

qRT-PCR reactions were performed in an iQ™5 Real-Time PCR Detection System (Bio-Rad Laboratories, München, Germany), by adding 10µl of iQ™ SYBR Green Supermix (Bio-Rad Laboratories), 4 µl of diluted cDNA, 0.5pmol of each primer, and water to a final volume of 20µl. After one initial incubation step at 95°C for 3min, amplifications were performed for 40 cycles with the following cycle profile: a denaturing step at 95°C for 15s followed by an annealing step at 60°C for 10s,

and an extension step of 72°C for 10s. Fluorescence data were collected during the 72°C step, and the specificity of qRT-PCR products was confirmed by performing a melting temperature analysis at temperatures ranging from 55°C to 95°C in intervals of 0.5°C, and running the PCR product in agarose gel to confirm the existence of a unique band of the expected size.

For all genes studied, three independent cDNA samples of each experimental condition were evaluated in technical duplicates, giving a total of 6 technical replicates for each treatment.

The raw, background-subtracted fluorescence data provided by the iQ5 software (version 2.0) were analyzed by the real-time PCR Miner software (version 2.2) (www.ewindup.info/miner/version2) (Zhao and Fernald, 2005). The resulting PCR efficiency and fractional cycle number of the threshold (CT) were used for transcript quantification. The efficiency for each gene was calculated doing the arithmetic mean of all efficiencies given by PCR Miner.

For the relative quantification of the genes of interest (GI) was used the Pfaffl method (Pfaffl, 2001). The expression of a gene is calculated based in the efficiency (E) value of the primer pair used in the amplification reaction and the CT value obtained:

$$Expression = (E + 1)^{-CT}$$

(Equation 3)

The relative quantification is calculated with the Pfaffl equation. The Pfaffl method allows the use of different PCR efficiencies for the GI and the reference gene (RG).

$$Relative\ quantification_{GI} = \frac{(E_{GI} + 1)^{-CT_{GI}}}{(E_{RG} + 1)^{-CT_{RG}}}$$

(Equation 4)

Relative quantification of gene expression of the different tissues/treatments (T/T) was normalized in relation to the arithmetic mean of the shoot control treatment (SC) relative quantification:

$$Relative\ quantification\ in\ relation\ to\ Shoot\ control = \frac{Relative\ quantification_{T/T}}{Mean\ Relative\ quantification_{SC}}$$

(Equation 5)

A Two-tailed t-test of significance (Microsoft Office Excel, 2007) was used to compare the 4 conditions between them in each tissue.

6.7. Northern Blot

6.7.1. Electrophoresis and Blotting of RNA

All RNA samples were separated by electrophoresis, in 15% acrylamide gels (15x30cm). The gel solution (50ml) is composed by: 25g urea, 5ml MOPS 10X (200mM, pH7.0), 18.75ml 40% acylamide:bisacrylamide (19:1) and water to 50ml. The solution can be warmed in the microwave or water bath to dissolve the urea, but before use, the solution may cool at room temperature (can be

stored at 4°C). Polymerization of acrylamide was achieved by adding to each 12ml of solution, 4µl TEMED and 50µl 10% APS. The gel was allowed to polymerize at least 1 hour and before loading the samples the gels have to be pre-run about 1 hour, at 110V, in 1X MOPS buffer.

The RNA samples and the small RNA molecular weight marker (New England Biolabs, Ipswich, UK) were prepared by joining 10µl of formamide loading buffer with 15µg of RNA or 0.25µl of molecular weight marker, at a final volume of 20µl. Before loading the samples were denatured at 65°C for 10min. After loading, the gels were run in a Mini-Protean II Electrophoresis Device (Bio-Rad Laboratories) with 1xMOPS buffer, at 100V, for 3 hours (until all the xylene cyanol dye runs out of the gel), stained with Sybr Safe (Invitrogen) and visualized in a UV image acquisition system (Gel Doc, Bio-Rad Laboratories) (Supplementary Figure 2).

6.7.2. Cross-Linking of RNA

The RNA was transferred to a Hybond-NX membrane (GE Healthcare (formerly Amersham Biosciences) Piscataway, New Jersey, USA) by electro blotting with Bio-Rad's Trans-Blot Semi-Dry System (Bio-Rad Laboratories), at 20V for 60min, 4°C. After this, the gels were stained again with Sybr Safe (Invitrogen) and visualized (Supplementary Figure 2). In this work the chemical cross-link was chosen instead of the UV cross-link because it greatly improves the detection of small RNA by northern blot hybridization (Pall et al, 2007) A solution consisting of 0.13M 1-methylimidazole, pH 8.0 and 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) to a final concentration of 0.16M was prepared immediately before the cross-link (Pall et al, 2007). The hybridization membrane was placed, with the RNA side faced up on a 3MM Whatman paper saturated with the cross-link solution wrapped in SARAN wrap and incubated at 60°C for one hour (Pall et al, 2007).

6.7.3. Hybridization

The membranes were pre-hybridized with the ULTRA Oligo Hybridization buffer (Ambion) for about 2 hours. The Locked Nucleic Acid (LNA)-modified oligonucleotide (Exiqon, Vedbaek, Denmark) complementary to miR168 and miR162 and the molecular weight marker probe were labeled simultaneously in the same reaction. The labeling reaction consisted of: 1pmol of LNA-modified oligonucleotide probes, 0.3µl of molecular weight maker probe (New England Biolabs), 4µl of Kinase buffer, 1µl of T4 polynucleotide Kinase (10U/µl) (Invitrogen) and 5µl of γ P³²-ATP (PerkinElmer, Waltham, Massachusetts, USA) and ddH₂O to a final volume of 10µl. After incubation incubated for 1 hour at 37°C, the unincorporated radioisotopes were removed with Illustra G-25 MicroSpin columns (GE Healthcare) and the purified labelled probe was added to the hybridization solution used to perform the pre-hybridization step. Hybridizations were carried over night at 42°C (Várallyay et al, 2007). The membranes were washed in 2x SSC (20x SSC stock solution: 3.0M sodium chloride, 0.3mM sodium citrate, pH7.0) more 0.1% SDS (Sodium Dodecyl Sulphate) at 42°C, wrapped in SARAN wrap without allowing the membranes to dry and finally exposed in a phosphorimager cassette for 16-48 hours (depending on the signal detected in the membranes with a Geiger counter). The images were obtained, using a PhosphorImager (Storm 860, GE Healthcare). When necessary the

membranes were striped with boiling 0.1% SDS. In the end the membranes were hybridized with the small nuclear RNA U6 prove used as loading control.

7. Part 1

7.1. Introduction

This work was initiated using *M. truncatula* plants subjected to 4 different conditions of water deficit status: the control plants (Ct), moderate water deficit (MWD), severe water deficit (SWD) and recovery (Rec), as described in the Material and Methods section. Leaf relative water content (RWC) has been used as an indicator of the plant water status (Chaves, 1991; Lawlor and Cornic, 2002) due to the close relationship found to exist between RWC and predawn water potential measured during progressive water deficit in various plant species (Cruz de Carvalho et al, 1998; Scotti Campos et al, 1999).

The RWC was used as one physiological indicator of the plants water status in the 4 water deficit conditions. But in this way we are just analyzing responses to water deficit at a physiological level and because of this we decided to determine if the plant has responses to water deficit at the cellular/molecular level in accordance to the RWC.

We decided to investigate the expression of some water deficit-related genes in the 4 water deficit conditions, to establish a relation between the plant relative water content (Physiological condition) and the level of transcription of water deficit regulated genes. For this, we selected MI-1-PS and Dehydrin that were tested in a previous work with water deficit in *M. truncatula* (Matilde et al, unpublished results) and ERD1 and PIP1;5 were selected for been reported to have differential expression in water deficit (Kiyosue et al, 1993; Alexandersson et al, 2005).

7.1.1. Myo-inositol-1-phosphate Synthase (MI-1-PS)

L-myo-Inositol-1-phosphate synthase (1L-MI(1)P₁) (MI-1-PS) does the conversion of D-Glucose-6-P to 1L- myo-Inositol-1-phosphate, that constitutes the first step of myo-Inositol (MI) biosynthesis. MI is a very important molecule with several cellular functions: substrate for auxin conjugates; oxidation to D-glucuronic acid that have a role in the biogenesis of uronosil and pentosyl units of pectins, hemicelluloses and related structures in plant cell walls; precursor of methylated inositols like pinitol and ononitol that are osmoprotectors (Loewus and Murthy, 2000).

7.1.2. Early Responsive to Dehydration 1 (ERD1)

The designation Early Responsive to Dehydration 1 (ERD1) gene was given because of its early response (1h) to water deficit in *A. thaliana*, increasing its expression (Kiyosue et al, 1993). ERD1 encodes a ClpC-like protein, which sequence analysis suggests that may interact with the chloroplast-localized ClpP protease, to facilitate ATP-dependent proteolysis (Kiyosue et al, 1993; Nakashima et al, 1997; Weaver et al, 1999). Is chloroplast targeted and present within the soluble fraction of the chloroplast, presumably in the stroma (Weaver et al, 1999).

7.1.3. Plasma membrane intrinsic protein 1;5 (PIP1;5)

Aquaporins are channel forming membrane proteins with the extraordinary ability to combine a high flux with a high specificity for water across cell membranes (Shao et al, 2008). They can be localized

in the plasma membrane, tonoplast, endoplasmic reticulum or intracellular vesicles (Maurel et al, 2008; Shao et al, 2008). On the basis of sequence homology, aquaporins in most plant species can be divided into four subgroups: (1) the plasma membrane intrinsic proteins (PIP) (with two phylogenetic subgroups, PIP1 and PIP2); (2) the tonoplast intrinsic proteins (TIP); (3) the nodulin-26-like intrinsic membrane proteins (NIP); (4) small basic intrinsic proteins (SIPs) (Johanson et al, 2001). PIP1–PIP2 interaction in the endoplasmic reticulum is required for in plants PIP1 trafficking to the plasma membrane to modulate plasma membrane permeability (Zelazny et al, 2007). Otherwise PIP1 will be retained in the endoplasmic reticulum, and PIP2 transported to the plasma membrane.

7.1.4. Dehydrin

Dehydrins is a group division (group 2 of 6) of the late embryogenesis abundant (LEA) proteins (Dure et al, 1989) and they are proteins with small molecular weight (10-30kDa) (Hong-Bo et al, 2005). LEA proteins are formed during the late period of seed development a phase accompanied by dehydration (Hong-Bo et al, 2005). It is interesting that appearance of the LEA proteins in embryony tissue depends on content of endogenous ABA, whose level is usually increased during seed formation and especially before their maturation (Allagulova et al, 2003). Dehydrins are considered stress proteins that protect the plant against dehydration (Allagulova et al, 2003) and their synthesis is induced ABA treatment (Hellwege et al, 2004).

7.2. Results

7.2.1. Selection of a reference gene for water deficit conditions in *M. truncatula*

For MI-1-PS, ERD1, PIP1;5 and Dehydrin transcript quantification we decided to use the technique of Quantitative Real-Time Polymerase Chain Reaction (qRT-PCR) that have the advantage of the need of much less RNA compared with the Northern Blot technique, from ng to µg respectively. The relative quantification is based on the relative expression of a target gene versus a reference gene. (Pfaffl et al, 2001) The reference(s) gene(s) have the function to control the pipeting errors and the RNA quantity used to make cDNA. For this, the gene expression has to be equal in all treatments, or in the case of more then one reference gene used, their mean has to be constant in all condition studied.

Six genes were tested to select the reference gene(s) for gene expression analysis in the shoots and roots of the plants in the 4 conditions. The genes selected were: 60S ribosomal protein L2 (RPL2C), Adenine Phosphoribosyl Transferase 1 (APRT), Actin 7 (ACT7), Actin 11 (ACT11), Elongation Factor 1-Alpha (ELF-1a) and Histone Deacetylase 3 (HDA3) (Supplementary Table 3 and Supplementary Table 4). These genes are usually used to do screens in plants and referred as house keeping genes for reference genes (Nicot et al, 2005), less in the case of HDA3 that was selected just because we had results indicating that it could have a stable expression in the conditions tested.

To perform the selection of the best reference gene(s) were used geNorm and NormFinder softwares. The first relies on the principle that the expression ratio of two ideal internal control genes is identical in all samples, regardless of the experimental condition or tissue type. In this way, variation of the expression ratios of two real-life housekeeping genes reflects the fact that one (or both) of the genes is (are) not constantly expressed, with increasing variation in ratio corresponding to decreasing

expression stability (Vandesompele et al, 2002). The second evaluate the overall expression variation and the variation across the sample subgroups (treatments) (Andersen et al, 2004). The data used in the softwares was the expression level of the genes tested (Equation 3, Methods) from shoots (Figure 6, B) and roots (Figure 6, C) individually and together (Figure 6, A).

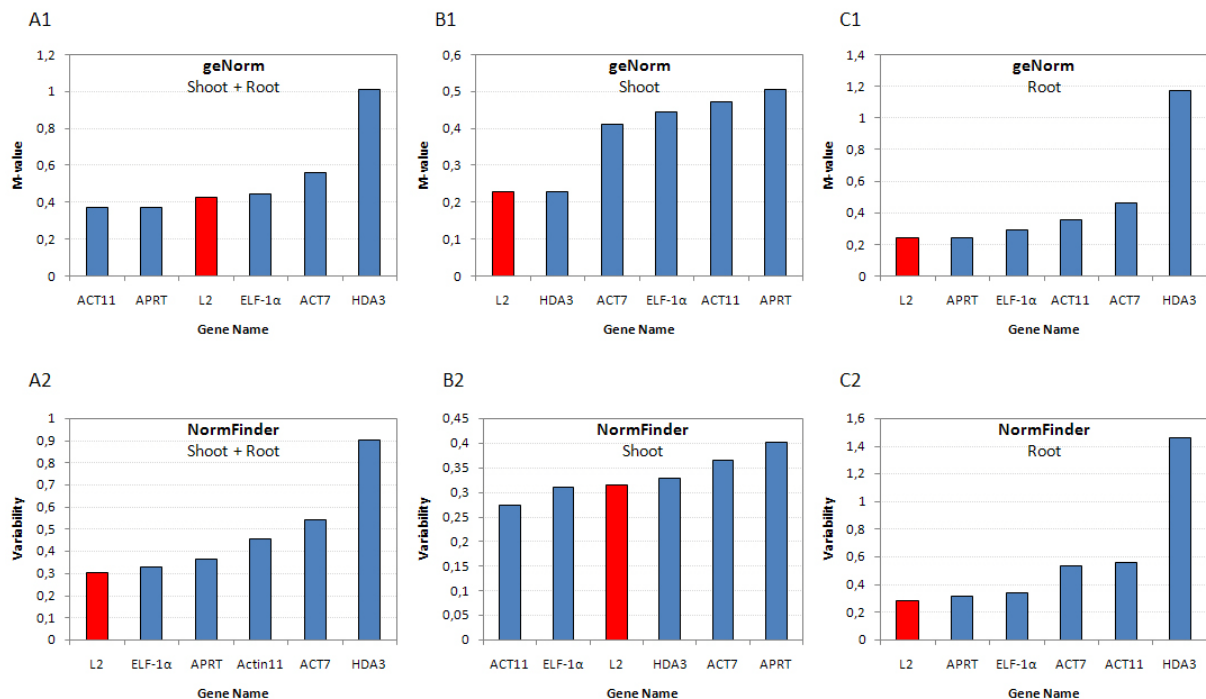


Figure 6: Selection of the best reference gene(s) for *M. truncatula* under water deficit. The expression values of the genes (Equation 3, Methods) were used in the geNorm (A1, B1, C1) and NormFinder (A2, B2, C2) softwares to determine the best gene. The analysis was performed with the expression levels calculated for the genes in shoot (B) and root (C) individually and together (A).

L2 was selected as the best reference gene (Figure 6), because it presented a more stable expression than the rest of the genes in all conditions tested. In the geNorm analysis, L2 was always in the group of the first 3 genes selected to have the most similar expression (Figure 6, A1, B1 and C1). But no other gene besides L2 stayed in the top 3, in this way is difficult to select more than one gene to use as reference genes. From this, we decided to select just one gene as reference gene. NormFinder selected L2 as the gene with the smallest variation between the 4 conditions for almost all analysis (Figure 6, A2 and C2) except in one of the situation (Figure 6, B2), but its difference from the best gene (ACT11) was very little. Usually normalization against a single reference gene is not acceptable (Bustin et al, 2009), but in our analysis L2 proved to be the best gene comparing to the rest of the genes.

7.2.2. The relation between the plant water status and the expression of water deficit responsive genes

In our water deficit conditions, the average RWC in Ct plants was around 80%. In water deficit treated plants the mean was 50% for MWD and 30% for SWD, and Rec plants had a RWC of about 30% after 8 days of water deprivation and then recovered to RWC of 80% after 3 days of watering. With the objective to establish a relation between the RWC and the expression of MI-1-PS, Dehydrin, ERD1

and PIP1;5 genes (Supplementary Table 5) we analyzed these genes expression in shoots and roots of *M. truncatula* plants subjected to the 4 water deficit conditions described above. The first two genes were identified as being regulated by water deprivation in *M. truncatula* in a previous study (Matilde et al, 2007, not published) and the other two were reported to have differential expression in water deficit in *A. thaliana* (Kiyosue et al, 1993; Alexandersson et al, 2005).

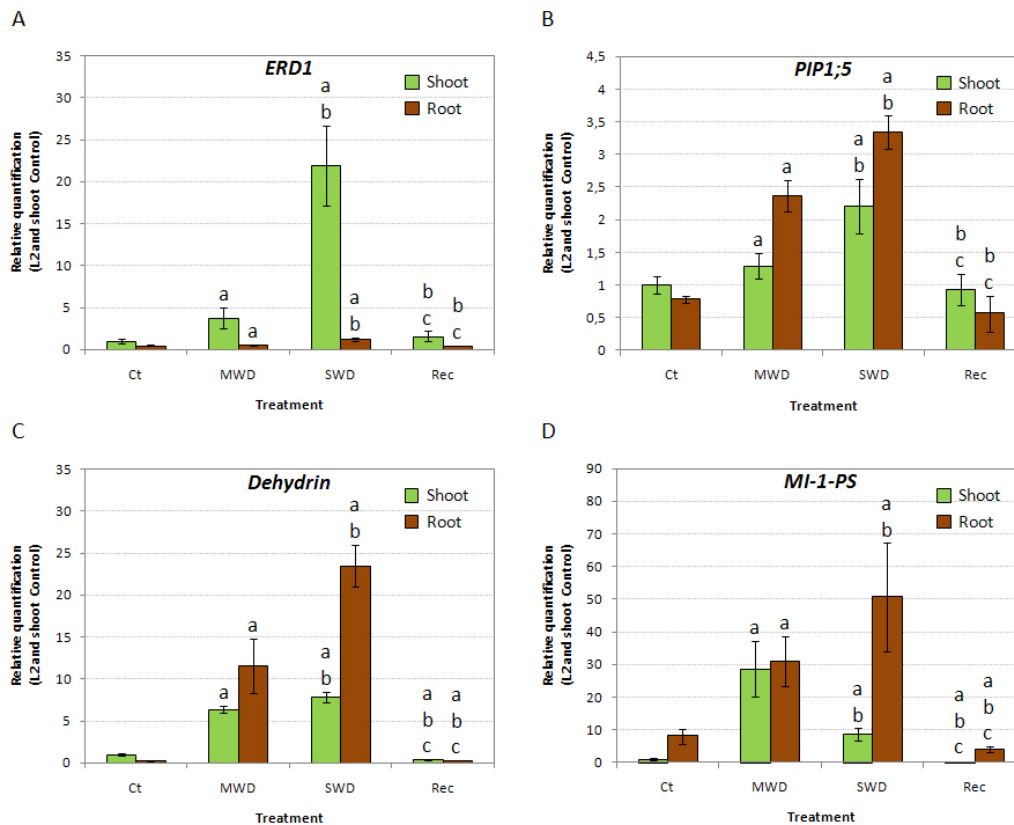


Figure 7: Relative expression levels of the genes ERD1, PIP1;5, MI-1-PS and Dehydrin determined by quantitative Real Time PCR. The shoot (green) and roots (brown) of *M. truncatula* are the tissues analyzed in the different water treatment conditions imposed to the plants. Values are the mean of two technical replicates of three independent cDNAs for each treatment (giving a total of 6 technical replicates) and bars represent standard errors. The relative expression was calculated using L2 as reference gene and normalized against to the shoot control treatment (Equation 5, Methods). A Two-tailed t-test of significance was applied, and was considered a significant difference with a p-value<0.05: a, significant difference with control of the same tissue, b, with MWD, and c, with SWD. (A) ERD1, Early Responsive to Dehydration 1, (B) PIP1;5, Plasma Membrane Intrinsic Protein 1;5, (C) Dehydrin, (D) MI-1-PS, L-Myo-inositol-1-Phosphate Synthase. Ct, Control; MWD, Moderate Water Deficit; SWD, Severe Water Deficit, Rec, Recovery.

All four genes increased their expression with the water deficit treatment in both tissues studied and returned to normal, control levels in the recovery treatment (Figure 7). Their expression was inversely related to the plants RWC, this is, at MWD (RWC of 50%) the genes increase their expression comparison to control (RWC of 80%) and at SWD (RWC of 30%) the expression levels are higher, returning to control levels in recovered plants (RWC of 80%). The only exception is MI-1-PS in the shoot (Figure 7, D) that in SWD decreases its expression relative to MWD.

7.3. Discussion

7.3.1. The regulation of MI-1-P, ERD1, PIP1;5 and Dehydrin in water deficit

The first objective of this work was to find a relation between the plant relative water content (physiological condition) and the level of transcription of water deficit regulated genes. All four genes presented a relation between the two variables, that is, when the RWC of the plant increased it was accompanied with the increase of those genes expression (Figure 7), except for MI-1-PS transcript levels in the shoots that decreased from MWD to SWD (Figure 7, D). Our observations seems to be in accordance with the levels of myo-inositol detected in leaves of *Nicotiana tabacum* plants subjected to water deficit conditions, in which myo-inositol concentration increased until day four without water (MWD corresponds to 5 days of suppression of soil irrigation, Methods) and then decreased along the water deficit treatment (SWD corresponds to 8 days of water deficit, Methods), but with values superior of those measured in day zero of the treatment (Sheveleva et al, 1997).

In *A. thaliana* it was shown that MI-1-PS reduced its expression in the presence of increasing concentrations of inositol (Johnson et al, 1995). In *Saccharomyces cerevisiae* was demonstrated that MI-1-PS regulation by inositols happens because their promoters have a conserved repeated element known as UAS_{INO}, which is the binding site for two heterodimeric transcription factors, INO2 and INO4 (Majumder et al, 1997). The decrease in the expression of MI-1-PS in SWD shoots in relation to MWD (Figure 2, D), was maybe due to the increase in the inositol concentration in the cells. The reason why in roots there is not one inhibition of the gene (Figure 7, D) is unknown. Myo-inositol transport through the xylem from the roots to the shoots increased in salt stressed plants (Nelson et al, 1998). We can hypothesize that in our plants inositol is transported from the roots to the shoots, so that the repression happens just in the shoots.

In our results ERD1 appeared to be strongly induced by water deficit in shoots as demonstrated by Nakashima and coworkers (1997) in *A. thaliana*. In stress conditions like this, chloroplast ClpP proteases could be important to degrade unassembled or misfolded proteins (Halperin and Adam, 1996; Majeran et al 2000). ERD1 is a ClpC-like protein, is suggested that may interact with the chloroplast-localized ClpP protease, to facilitate ATP-dependent proteolysis (Kiyosue et al, 1993; Nakashima et al, 1997; Weaver et al, 1999), so if its expression increased in plants under water deficit conditions it could possible mean that in the chloroplasts the number of unassembled or misfolded proteins increased because of the stress condition imposed to the plants.

The aquaporin here studied is homologous of PIP1;5 from *A. thaliana*. In water deficit conditions the expression of several aquaporins could be up-regulated, down-regulated or kept its expression levels constant (Alexandersson et al, 2005). Although the four classes are conserved among all plant species, the aquaporin gene family shows signs of rapid and recent evolution and orthologs can not necessarily be distinguished between species (Maurel et al, 2008). MtPIP1;5 has increased its expression in *M. truncatula* but in *A. thaliana* the expression level of AtPIP1;5 reduced its expression (Alexandersson et al, 2005). This could be explained because the two genes got different functions and different regulatory regions in the evolution of both species.

8. Part 2

8.1. Introduction

miRNA biogenesis is under feedback regulation such that two key players in miRNA biogenesis and function are themselves regulated by miRNAs. DCL1 mRNA has a binding site for miR162, which leads to the cleavage of DCL1 mRNA (Xie et al, 2003). Consistent with this, DCL1 mRNA levels are elevated in the *hen1-1* mutant, in which the abundance of miR162 is reduced. miR168 has a binding site in AGO1 mRNA and leads to AGO1-mediated cleavage of AGO1 mRNA (Vaucheret et al, 2004). Vaucheret and coworkers (2006) described the transcriptional co-regulation in time and space of miR168 and AGO1, the miRNA is regulated by its target mRNA/protein levels, in a way that, if the expression of AGO1 increases, then it will load more miR168, that in turn will silence more AGO1 mRNA, returning it to normal level. The model proposed imply that a perturbation in the level of miR168, AGO1 mRNA, or AGO1 protein would induce compensatory adjustments to bring AGO1 levels back to equilibrium. They proposed that the miR168 duplex loading to AGO1 is very inefficient, and because of this the slight alteration of AGO1 level will induce a feedback response to reestablish the usual levels. AGO1 is silenced by cleavage (Vaucheret et al, 2004) and by translational repression that has a stronger silencing effect, of three fold compared with the cleavage situation (Lanet et al, 2009).

The second objective of this work was to study the feedback regulation of miRNA biogenesis in *M. truncatula* under water deficit conditions. To achieve this goal, we need to study the relation between, miR162 and miR168 expression levels and their respective targets, DCL1 and AGO1.

8.2. Results

8.2.1. Identification of *M. truncatula* miR168

Most mature miRNAs are evolutionary conserved from species to species within the Plant Kingdom, which provides an easy way to predict new miRNA homologous in most plant species (Dezulian et al, 2005; Zhang et al, 2006a). The sequence of miR162 from *M. truncatula* was downloaded from the mirBase (Dezulian et al, 2005), but the sequence of miR168 was not available in the database (until July of 2009). Szittyá and coworkers (2008) sequenced a small RNA library and identified the mature sequence of miR168, but did not publish its sequence. In a later work, the mature sequence of miR168 was published, but the secondary structure of its precursor and its loci were not determined (Jagadeeswaran et al, 2009). Thus, we decided to perform a bioinformatic search for miR168 in the *M. truncatula* genome (Mt2.0), taking into account the following criteria for its annotation (Ambros et al, 2008): expression criteria - consists in distinction of a RNA with the size of the mature form of the miRNA by northern blot; biogenesis criteria – bioinformatic homology search using as query the sequence of miR168 from *G. max*, prediction of its secondary structure, determination of MFE (Ambros et al, 2008), AMFE, MFEI and A + U content (Zhang et al, 2006b).

The *M. truncatula* miR168 mature and precursor sequences were predicted using *G. max* miR168 and pre-miR168. As shown in Table 1, there are 3 loci for miR168 in the *M. truncatula* genome (Mt2.0), all

located on chromosome 5 (Figure 8), with two mature miR168 sequences having the same nucleotide sequence and a third one with two mismatches.

Table 1: Characteristics of the mature and precursor sequences of the putative miR168 sequences of *M. truncatula*. L. (bp), Length in base pairs; N.M., Nucleotide mismatches; MFE, minimum free energy; AMFE, adjusted MFE; MFEI, minimum folding free energy index; BAC AC, BAC accession number.

miRNA	Mature miRNA			pre-miRNA + 18bp 5' and 3'						Genome
	mature miRNA sequence	RNA Blot	L. (bp)	N.M.	Arm	A+U (%)	MFE (ΔG)	AMFE (ΔG)	MFEI	BAC AC
miR168a	UCGCUUGGUGCAGGUCGGGAA	Yes	21	0	5'	44.4	-68.8	-55.5	1.00	CU302345
miR168b	UCGCUUGGUGCAGGUCGGGAA	Yes	21	0	5'	52.5	-53.5	-44.6	0.94	CU302337
miR168c	UUGCUUGGUGCUGGUCGGGAA	No	21	2	5'	53.9	-50.1	-39.2	0.85	AC126012

The predicted precursor miR168 sequences of *M. truncatula* have low MFE values, the AMFEs values are between -39,2 and -55,5 and the MFEIs are equal or higher than 0.85 and the A+U content between 44.4 and 53.9. These results are in conformity with Zhang and collaborators (2006b) and are good indicators of the presence of a miRNA gene sequence.

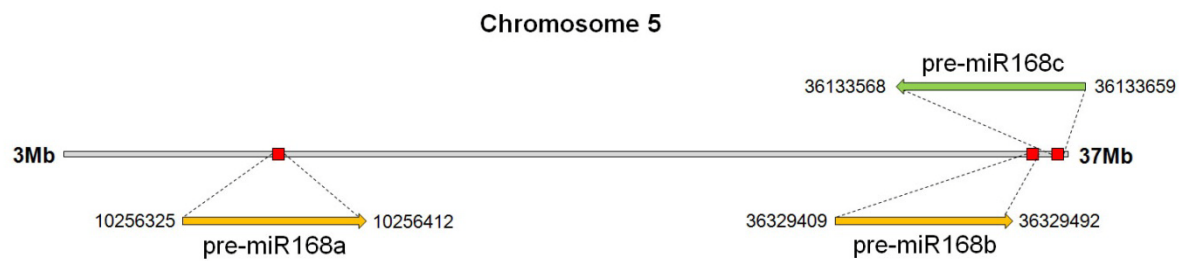


Figure 8: Localization in *M. truncatula* chromosome 5 of the precursor sequences of miR168a, b and c genes. The arrow is pointing from 5' end to 3' end of the pre-miR168 sequences.

The secondary structure of the pre-miRNAs was determined (Figure 9), and the mature sequence of miR168c presented the higher number of mismatches (5) with the miRNA* region. The region of complementarity of the mature miRNA with its miRNA* in the pre-miRNA presented the highest stability (less number of mismatches, wobbles or bulges) (Figure 9, positional entropy), comparing to other regions of the precursor.

The miRNAs in plants have high complementarity with their targets, making easy their search (Rhoades et al, 2002). When conservation of miRNAs among different species exists, follows the conservation of the targets of that miRNAs (the same target genes or genes that belong to the same family). (Jones-Rhoades and Bartel, 2004; Lu et al, 2005; Zhang et al, 2006a; Jagadeeswaran et al, 2009) The targets predicted for miR162, miR168a, b and c are shown in Supplementary Table 6 and Supplementary Figure 3. For miR162 the target gene belonged to the family of the Dicer-like proteins (DCL1) and for miR168a, miR168b and miR168c the target was the Argonaute 1 gene.

These targets were previously predicted in other plant species (Jones-Rhoades and Bartel, 2004; Lu et al, 2005) and their validation was done by 5' RACE (Rapid Amplification of cDNA Ends) in Jagadeeswaran et al (2009). With the modified 5'RACE technique, one of the mRNA target strands originated from miRNA directed cleavage is sequenced to confirm the identity and position of the cleavage of the cleaved target (Llave et al, 2002).

Although we could find three miR168 loci in the *M. truncatula* genome, we decided to continue this work with miR168a/b and miR162 and drop miR168c, because the data on this miRNA is less reliable than for the other two miR168. Also, Jagadeeswaran et al (2009) and Szittyá et al (2008) could not sequence miR168c, although these last authors describe sequences of miR168 with two mismatches in their small RNA libraries but at a very low number.

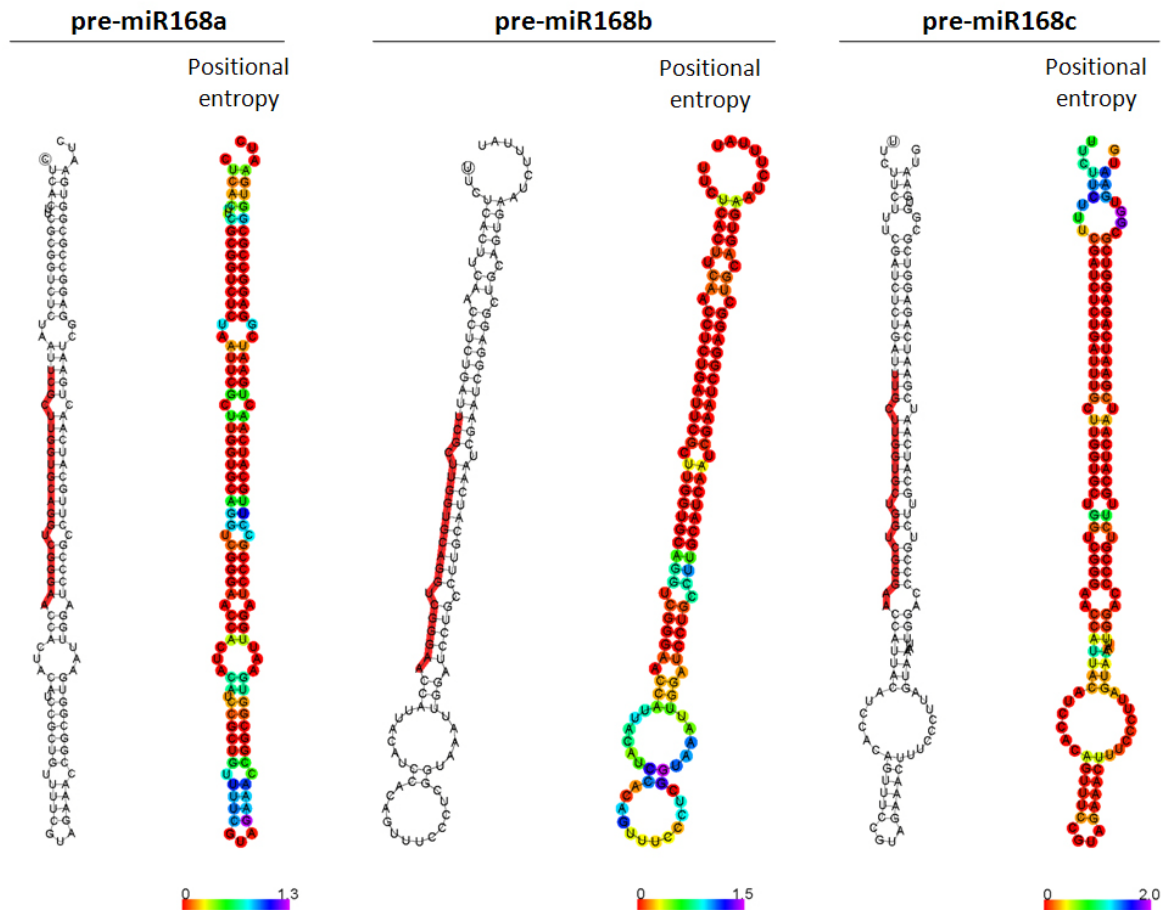


Figure 9: Secondary structure of *M. truncatula* miR168 precursors. The sequence of miR168a, miR168b and miR168c are evidenced in red at the secondary structure of the respective precursors (RNA folding/annotation software). The positional entropy (RNAfold) of each nucleotide in the precursor's secondary structure is shown.

8.2.2. The three loci of miR168 in *M. truncatula* and their evolutionary relationship

Some MIRNAs genes might have arisen from duplication of their target loci (Allen et al, 2004; Rajagopalan et al, 2006) or from MIRNAs (Gazzani et al, 2009) In *M. truncatula* there are 3 loci for MIR168 gene family in chromosome 5 (Figure 8), two in plus strand and the other in the minus strand. And doing a phylogenetic analysis with the precursor sequences (Figure 10) we stated that the first duplication event is between MIR168a and MIR168b/c, and the second is between MIR168b and MIR168c.

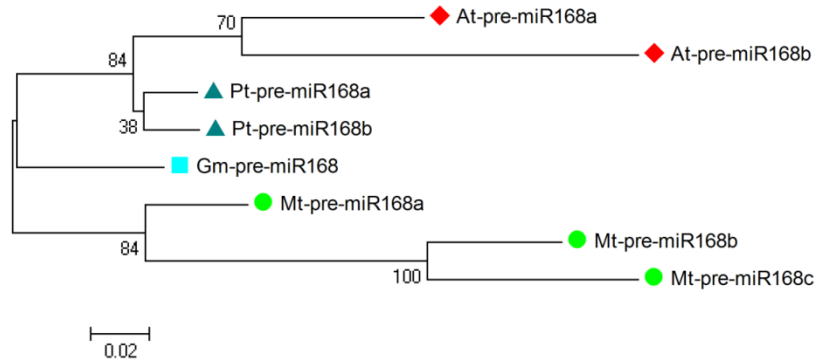


Figure 10: Phylogenetic tree of the pre-miR168 sequences from *M. truncatula* (Mt), *G. max* (Gm), *A. thaliana* (At) and *Populus trichocarpa* (Pt). For alignment was used the RCOFFEE software, that take into account the secondary structure of the RNA used for the alignment. A Neighbour joining tree was constructed using the MEGA 4.0 software.

8.2.3. Expression analysis of miR162 and miR168a/b in *M. truncatula*

Northern blot analysis of miR162 and miR168a/b expression in different plant parts of *M. truncatula* (Shoot, Root, Seedling, Seedpod and Flower) was very similar between both of them (Figure 11). Their expression was higher in the roots and lower in the seedpods and flowers. For miR162, two bands were visible (Figure 11): one band of 21bp that correspond to the miR162 size (Table 1) (Szitty et al, 2008) and an extra less strong band of 24bp. For miR168a/b, again two bands are visible, one of 21bp that corresponds to the miRNA and a faint band of 24bp. The reason for these extra bands is that DCL3 competes with DCL1 for the same miRNA precursors to produce small RNAs molecules with 24nt (Vazquez et al, 2008). Nevertheless, the presence of a band of the correct size in all plant parts fills the expression criteria that distinct a RNA with the size of the mature form of the miRNA, indicating that both miR162 and miR168a/b are expressed in *Medicago truncatula*.

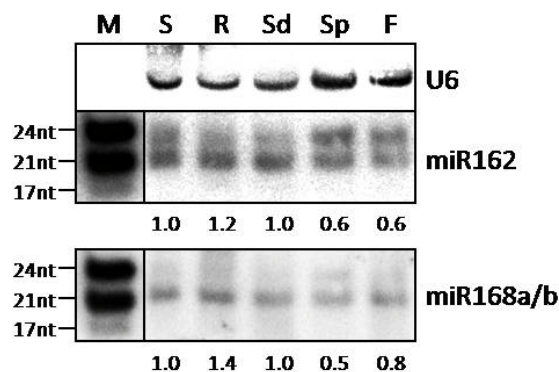


Figure 11: Northern blot of miR162 and miR168a/b in *M. truncatula* tissues. RNA gel blots containing total RNAs (15µg) from different samples of shoot (S), root (R), seedling (Sd), seedpod (Sp) and flower (F). They were hybridized with radioactive Locked Nucleic Acid (LNA) probes complementary to the miRNAs. The small nuclear RNA U6 was used as internal loading control for quantification of RNA gel blot signals (indicated down each lane), which were normalized against to the shoot samples. The membrane was first hybridized with miR168a/b prove and then stripped and rehybridized with miR162 prove. A molecular marker was used and presents three different sizes: 17nt, 21nt and 24nt.

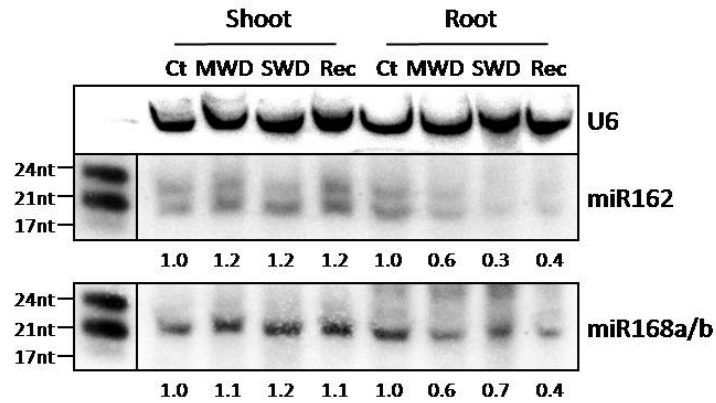


Figure 12: Northern blot of miR162 and miR168a/b of RNA samples of shoots and roots from *M. truncatula* plants with different water status. RNA gel blots contained 15µg of total RNA, for each sample. They were hybridized with radioactive Locked Nucleic Acid (LNA) probes complementary to the miRNAs. The U6 small nuclear RNA U6 was used as internal loading control for quantification of RNA gel blot signals (indicated down each lane), which were normalized against to the control condition of the shoots and roots samples. The membrane was first hybridized with miR162 probe and then striped and rehybridized with miR168a/b probe. A molecular marker was used and presents three different sizes: 17nt, 21nt and 24nt. Ct, Control; MWD, Moderate Water Deficit; SWD, Severe Water Deficit; Rec, Recovery.

8.2.4. Expression of miR162 and miR168a/b and their targets during water deficit

From the moment that the mature sequence of miR168 and miR162 and their targets were established in *M. truncatula* we could proceed for to their expression analysis in the water deficit conditions previously described.

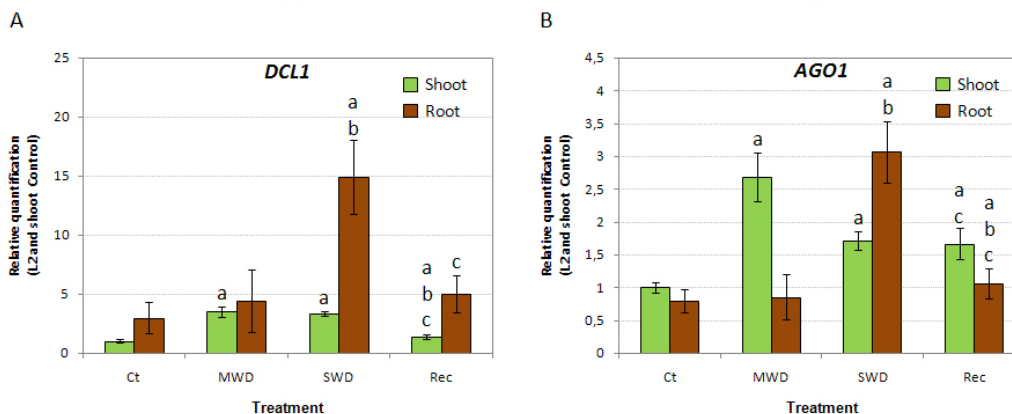


Figure 13: Relative quantification of miR162 and miR168a/b targets, Dicer-like 1 (DCL1) (A) and Argonaute 1 (AGO1) (B) mRNAs respectively, by quantitative Real Time PCR. The shoot (green) and roots (brown) of *M. truncatula* are the tissues analyzed in the different water treatment conditions imposed to the plants. Values are the mean of two technical replicates of three independent cDNAs for each treatment (giving a total of 6 technical replicates) and bars represent standard errors. The relative expression was calculated using L2 as reference gene and normalized against to the shoot control treatment (Equation 5, Methods). The AGO1 and DCL1 primer pair was designed to give one amplicon with the cleavage site of their corresponding miRNA. A Two-tailed t-test of significance was applied, and was considered a significant difference with a p-value<0.05: a, significant difference with control of the same tissue, b, with MWD, and c, with SWD. Ct, Control; MWD, Moderate Water Deficit; SWD, Severe Water Deficit, Rec, Recovery.

In the shoots, miR168a/b presented like miR162 a small increase in the water deficit treatments and recovery (Figure 12). The miR162 expression levels in the roots reduced from the control to MWD and

SWD (with the lowest value) and increased a little in the Rec treatment compared to SWD. miR168a/b reduced its expression in the roots from the control to MWD then the expression increased a little in SWD and reduced again in the recovery condition (with the lowest value) (Figure 12).

The expression levels of miR162 (DCL1) and miR168a/b (AGO1) targets (Supplementary Table 6) were quantified by qRT-PCR, to determine if there is a negative relation between the expression levels of the miRNAs and their respective targets. In all water deficit treatments and tissues analyzed there was an increase in the expression levels of DCL1 and AGO1 (Figure 13). The roots showed the highest increase between control and SWD, of about 5 and 3.5 fold, in DCL1 and AGO1 respectively (Figure 13).

8.3. Discussion

8.3.1. miR168 annotation

Three loci that correspond to putative genes of miR168 were found (Figure 8). The secondary structures of the precursors were evaluated and are in agreement with the biogenic criteria, but the expression criterion was just validated for miR168a/b by northern blot (Figure 11 and 12). Jagadeeswaran et al (2009) and Szittyá et al (2008) sequenced miR168a/b, but Szittyá et al (2008) found miR168 mature sequence with 2 mismatches with the miR168a/b. But the proportion of reads with two mismatches was approximately 4%, and in the library sequences with two mismatches were identified for almost all miRNAs, some with higher proportions. The only way to verify if miR168c is a miRNA is at least to do its blot (to verify the expression criteria) or sequence it with its miRNA* (Kim et al, 2009). We could not confirm if both precursors (pre-miR168a and pre-miR168b) are responsible for the mature miR168a/b that appears in the blot. One way to confirm this is to perform a promoter analysis of the MIR168a and MIR168b genes, to know if both genes are expressed and present variations between them.

8.3.2. MIR168a/b /c genes phylogenetic analysis

A. thaliana and *P. trichocarpa* have two genes for MIR168 in the genome (Figure 10) and Gazzani and coworkers (2009) suggested that these copies had their origin from two independent duplication events. Then *A. thaliana*, *P. trichocarpa* and *M. truncatula* share a common ancestral that had just one copy of the MIR168 gene and that by independent evolution in the each species lineage had occurred at least one duplication event (paralogous sequences) (Figure 10).

8.3.3. Regulation of miRNA biogenesis in water deficit by miR162

DCL1 is subjected to negative feedback regulation by miR162 (Reinhart et al, 2002; Xie et al, 2003) that maintain DCL1 at concentrations functionally sufficient, but not at limiting or excessive levels. In our results, miR162 is less expressed at the roots in water deficit (Figure 12), having the highest decrease in SWD. Correlating this with the levels of DCL1 expression (Figure 13, A), we have strong evidence that the regulation of the DCL1 gene is being done by miR162 in water deficit conditions. The possible implication of the DCL1 mRNA up-regulation is the increase on DCL1 protein levels. In this way, the production of mature miRNAs will increase. One possibility for this is the increase of the

exoribonucleases that degrade the miRNAs in water deficit (Ramachandran et al, 2008). In this way to compensate the degradation, the cell has to increase the production of mature miRNAs (up-regulation the DCL1 levels). To confirm this, the expression analysis of this exoribonucleases should be performed.

8.3.4. Regulation of miR168 and its target in water deficit conditions

Argonaute1 is targeted by miR168a/b (Rhoades et al, 2002) and this is needed for proper plant development, illustrating the importance of Argonaute 1 feed-back control by this miRNA (Vaucheret et al, 2004). In our results we observe in roots that miR168a/b decrease its expression in MWD, SWD and Rec (Figure 12) with the lowest expression in Rec, but AGO1 just increase its expression in SWD (Figure 13, B). These results are not completely coherent. To complete the analysis of miR168 and its regulation we would need to do a western blot in the shoots and roots in the different treatments, to determine the real implications of water deficit in the levels of AGO1 protein.

9. Part 3

9.1. Introduction

After analysis of the expression of miR162 and miR168a/b and their targets (DCL1 and AGO1) in water deficit, we decided to find other Dicer-like and Argonaute coding genes in *M. truncatula* genome and do the gene expression quantification of some of them to add more information to our study on the behavior of components of other small RNA biogenesis pathways during water deficit.

9.1.1. Plant Dicer-like proteins

Plants are characterized for having at least four DCLs (Margis et al, 2006), with different roles: DCL1 generates miRNAs with ~21nts; DCL2 produce siRNAs with 22nts, associated with virus defense and production of natural antisense-siRNAs (nat -siRNAs) from natural cis-acting antisense transcripts; DCL3 generates siRNAs with 24nts, that guide chromatin modification; DCL4 generates trans-acting siRNAs with 21nts (Vaucheret et al, 2008; Liu et al, 2009).

Plant DCL proteins contain six types of domains (Margis et al, 2006):

- **DEAD-helicase box:** kinetic analysis of wild-type and mutant human Dicer proteins showed that the DEXD/H-box domain might have one auto-inhibitory function, because removal of this domain increases the cleavage rate. This suggests that before catalysis can occur, the DEXD/H-box domain imposes on the protein a non-productive conformation that must be arranged, perhaps upon assembly with its molecular partners (Jinek and Doudna, 2009);
- **Helicase-C:** with unknown function;
- **DUF283:** DUF283 has a fold similar to the double-stranded RNA-binding domain (dsRBD) (Dlakić, 2006).
- **PAZ:** binds to single- and double-stranded RNAs at the 3' overhangs (Olmedo and Guzmán, 2008).
- **RNase III:** this domain is present in two copies (RNase IIIa and RNase IIIb) in DCLs. The two RNase II domains form an intramolecular dimer and the active site of each domain is involved in catalytic activity in dsRNA cleavage. The distance between the 3'-overhang-binding pocket of the PAZ domain and the active site of the RNAase IIIa domain is 65 Å, and matches the length of the siRNAs produced by it (Jinek and Doudna, 2009).
- **dsRBD:** plant DCLs have usually two copies of this domain, but in the case of DCL2 there is only one. dsRBD domain not only bind to dsRNA but also function as protein-protein interaction domain, interaction specially with the HYL1 protein (Margis et al, 2006).

9.1.2. Plant Argonaute proteins

AGOs are large proteins (90–100 kDa) and multidomain proteins that contain one variable N-terminal domain and conserved C-terminal PAZ, middle (MID) and PIWI domains (Tolia and Joshua-Tor, 2007; Hutvagner and Simard, 2008; Jinek and Doudna, 2009). The crystal structure of a full-length Argonaute protein, from the archaeal species *Pyrococcus furiosus*, showed that the sequence motif

originally defined as PIWI domain by Lorenzo Cerutti et al (2000) consists of two structural domains, termed MID and PIWI. The Argonaute domains have specific functions:

- **PAZ:** is a RNA-binding domain that specifically recognizes the 3' ends of ssRNAs suggesting that it might function as a module for anchoring the 3' end of the guide RNA (Vaucheret, 2008; Jinek and Doudna, 2009).
- **MID:** the 5' phosphate group of the guide RNA strand is buried in a deep pocket at interface between the MID domain and PIWI domain (Jinek and Doudna, 2009).
- **Piwi:** this domain adopts a fold similar to that in RNaseH, an endoribonuclease that cleaves RNA-DNA hybrids, suggesting that Piwi domain is responsible for the "slicer" activity of Argonaute protein in the RISC complex (Jinek and Doudna, 2009). The cleavage activity is carried out by an active site on the PIWI domain usually presenting an Asp-Asp-His (DDH) motif (Rivas et al, 2005; Bumberger and Baulcombe, 2005).

In plants distinct types of sRNAs associate with different Argonaute proteins on the basis of the identity of the 5' nucleotide. AGO1 binds mainly to RNAs with a uridine at their 5' end, whereas AGO2 and AGO4 recruits RNAs with a 5' adenosine and the AGO5 predominantly bind sRNAs that initiate with a cytosine (Mi et al, 2008).

For Argonaute to exert its slicer activity the perfect complementarity around the cleavage site in the guide-target duplex is required. The base-pairing around the 10-11-base ensures correct orientation of the active site for the cleavage (Jinek and Doudna, 2009).

9.2. Results

9.2.1. Molecular characterization of other AGOs and DCLs in *M. truncatula*

Blastn and tBlastn search in *M. truncatula* genome databases using *A. thaliana* AGOs and DCLs cDNA and proteins sequences (Supplementary Table 1) resulted in the identification of 3 sequences encoding Dicer-like genes (MtDCLs) and 9 sequences for Argonautes (MtAGOs) (Table 2) (note that DCL1 and AGO1 were previously identified as miR168a/b and miR162 targets). MtAGO2 was just found in *M. truncatula* cDNA database MTGI, and MtDCL2, MtDCL3 and MtAGO11 were just found in *M. truncatula* annotated genome (Mt2.0) (Table 2).

The MtAGO7 genome sequence (CU179907_3) and CDNA sequence (TC114668) originated incomplete Argonaute protein sequences (Supplementary Table 7 and Supplementary Figure 4) when analyzed using the Conserved Domains Database (CDD) from NCBI for domain search. But both sequences complemented each other, then we decided to join both them in one (Supplementary Figure 4 and Supplementary Figure 6). The BAC sequence close to the MtAGO7 gene presents some regions with N entries, and this means that the sequencing on that region was of poor quality what may explain why the complete genomic region of the gene is not in the *M. truncatula* databases. With MtAGO11 a similar case happened but with two annotated incomplete AGO protein sequences (CT030192_30 and CT030192_31) (Supplementary Table 8 and Supplementary Figure 5) whose genes were spaced in the genome just by 282nt from each other (intron size range from 0.1 to 0.6 kb in plants, Carels and Bernardi, 2000). After applying the Fgenesh-H software (for prediction of multiple (alternative splicing) variants of potential genes in genomic DNA)(<http://linux1.softberry.com>) in the

genomic region that contains the two annotated genes, one coding sequence was obtained which translated protein gave one complete AGO protein (Supplementary Figure 5 and Supplementary Figure 7).

Table 2: Characteristics of Argonaute and Dicer-like genes and proteins identified in *M. truncatula* (Mt). Mt2.0 genecall, *Medicago truncatula* Genome Sequencing Project release version 2.0 (Mt2.0) accession number of the gene; ORF, Open Reading Frame; MTGI, *Medicago truncatula* Gene Index (MTGI) 9.0 reference; MW, Molecular Weight; pI, Isoelectric point; Chrom., Chromosome; NCBI Blastn and Blastp in *A. thaliana* database.

Gene Name	Mt2.0 Genecall	ORF Size (bp)	MTGI	Reading Frame	Protein			Chrom.	Genomic Region	Blastn Blastp
					Size (a.a.)	MW (KDa)	pI (pH)			
DCLs										
MtDCL1	AC150443_32	5820	TC129362		1940	218.32	6.22	7	31913677-31923713	DCL1
MtDCL2	AC192958_13	4119			1373	156.11	7.18	2	27599809-27590829	DCL2
MtDCL3	AC137830_51	5184			1728	192.35	6.68	3	19558574-19571225	DCL3
AGOs										
MtAGO1			TC126820	2	1040	115.42	9.47			AtAGO1
MtAGO12b	AC136450_38	1521	NP7267711		506	57.69	10.03	2	14863531-14868049	AtAGO10
MtAGO12a	AC160838_11	2631			876	85.97	9.00	8	25391911-25397439	AtAGO10
MtAGO2			TC135942	1	720	85.35	9.01			AtAGO2
MtAGO7	CU179907_3	3051	AW693202 BI309506		1016	116.65	9.44	5	17162372-17166459	AtAGO7
MtAGO4a	AC147429_4		TC114668	1	824	92.40	8.24	0		AtAGO4
MtAGO4b	AC131455_15	2726	TC114471		908	101.37	9.13	5	31663907-31669806	AtAGO4
MtAGO4c	AC131455_31	2607	TC112620		868	97.83	9.54	5	31672438-31678208	AtAGO4
MtAGO11	CT030192_30 CT030192_31	2661			886	101.10	9.18	3	1284064-1277890	AtAGO4

DCL genes encoded a protein with a molecular weight between 156.11 and 218.32 KDa, with a neutral isoelectric point ranging from 6.22 to 7.18 (Table 2) in accordance to Kapoor et al (2008). AGO proteins presented a lower molecular size ~100KDa and a basic isoelectric point between 8.24 and 10.03 (Table 2) (Kapoor et al, 2008).

9.2.2. *M. truncatula* AGOs and DCLs protein domains

To designate the *M. truncatula* AGOs and DCL genes found a neighbor joining phylogenetic tree was built containing the *M. truncatula* and *A. thaliana* DCLs and AGOs protein sequences (Figure 14, A and Figure 15, A). The DCLs and AGOs clustered into 3 subgroups, similar to those described by Margis et al (2006), Liu et al (2009) and Vaucheret (2008). The names were given in respect to their phylogenetic relationship with *A. thaliana* proteins (Figure 14 and Figure 15).

The search for protein domains using CDD from NCBI revealed the presence of DEXD, Helicase-c, DUF283, PAZ, RNaseIIIa/b and dsRBa/b in DCLs protein sequences (Figure 14, B and Supplementary Table 9). MtDCL2 has just one domain dsRB characteristic of DCL2 proteins in *A. thaliana*, *Oryza sativa* and *P. trichocarpa* (Margis et al, 2006). All AGO protein presented the domains DUF1785, PAZ and Piwi (Figure 15, B and Supplementary Table 10). The MtAGO12b has an incomplete Piwi domain and this is because this gene was sequenced in the end of the BAC and was no more BACs sequenced (July of 2009 in Mt2.0) that could provide a consensus sequence in the terminal region where the Argonaute is located.

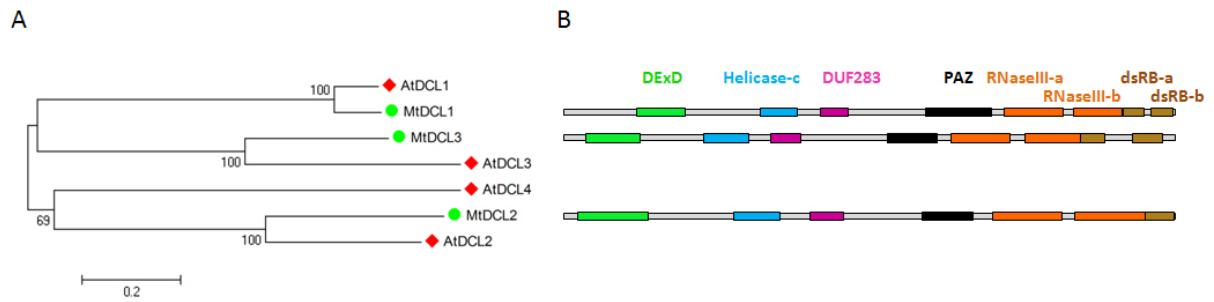


Figure 14: Evolutionary relationship of *M. truncatula* (Mt) and *A. thaliana* (At) DCLs (A) and characterization of the protein domains (B). The protein sequences were aligned using the CLUSTALW and a Neighbour joining tree was constructed using the MEGA4.0 software. The protein domains were obtained using the Conserved Domains Database (CDD) database of NCBI. The Dicer-like protein domains DEXD (green), Helicase-c in (blue), DUF283 (pink), PAZ (black), RNAase III (orange), dsRB (brown) are shown.

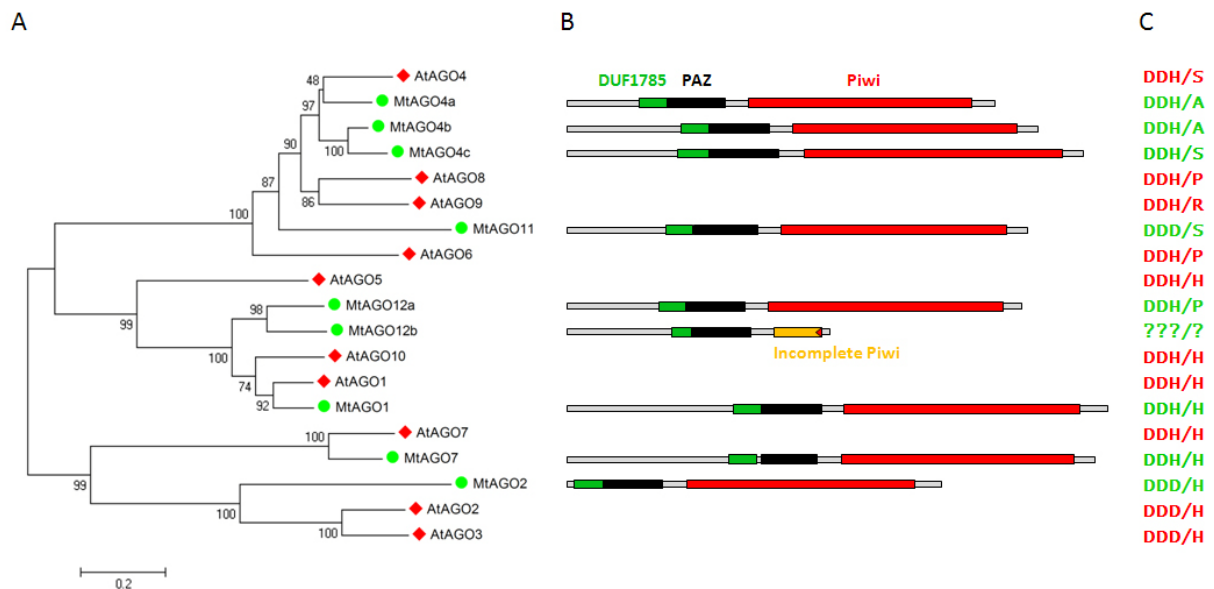


Figure 15: Evolutionary relationship of *M. truncatula* (Mt) and *A. thaliana* (At) AGOSs (A), characterization of the protein domains (B) and their catalytic center (C). The protein sequences were aligned using the CLUSTALW and a Neighbour joining tree was constructed using the MEGA4.0 software. The protein domains were obtained using the Conserved Domains Database (CDD) database of NCBI. The Argonaute protein domains DUF1785 (green), PAZ (black), Piwi (red) are shown. The catalytic center was obtained from the alignment of the Piwi domains from the AGO proteins, corresponding to the positions of the aspartate, aspartate and histidine (DDH) motif and the Argonaute 1 histidine at position 800 (H800). D, aspartate, H, histidine, S, serine, A, alanine, P, proline, R, arginine.

Structural studies in other organisms have shown that the Piwi domain of Argonaute proteins folds similar to RNaseH proteins (Song et al, 2004). Consistent with this observation, some Argonaute proteins in both plants and animals are known to cleave the target RNAs that have sequence complementary to the small RNAs (Liu et al, 2004; Baumberger et al, 2005).

The catalytic center of these catalytic proteins are known to possess three conserved metal chelating residues in the Piwi domain i.e. aspartate, aspartate and histidine (DDH), that function as the catalytic triad. In Arabidopsis AGO1, a conserved histidine at position 798 (H798) was also observed to be

critical for the endonuclease activity of AGO1 in vitro (Baumberger et al, 2005). We used a protein sequence from other database and because of that the histidine position passed from 798 to 800.

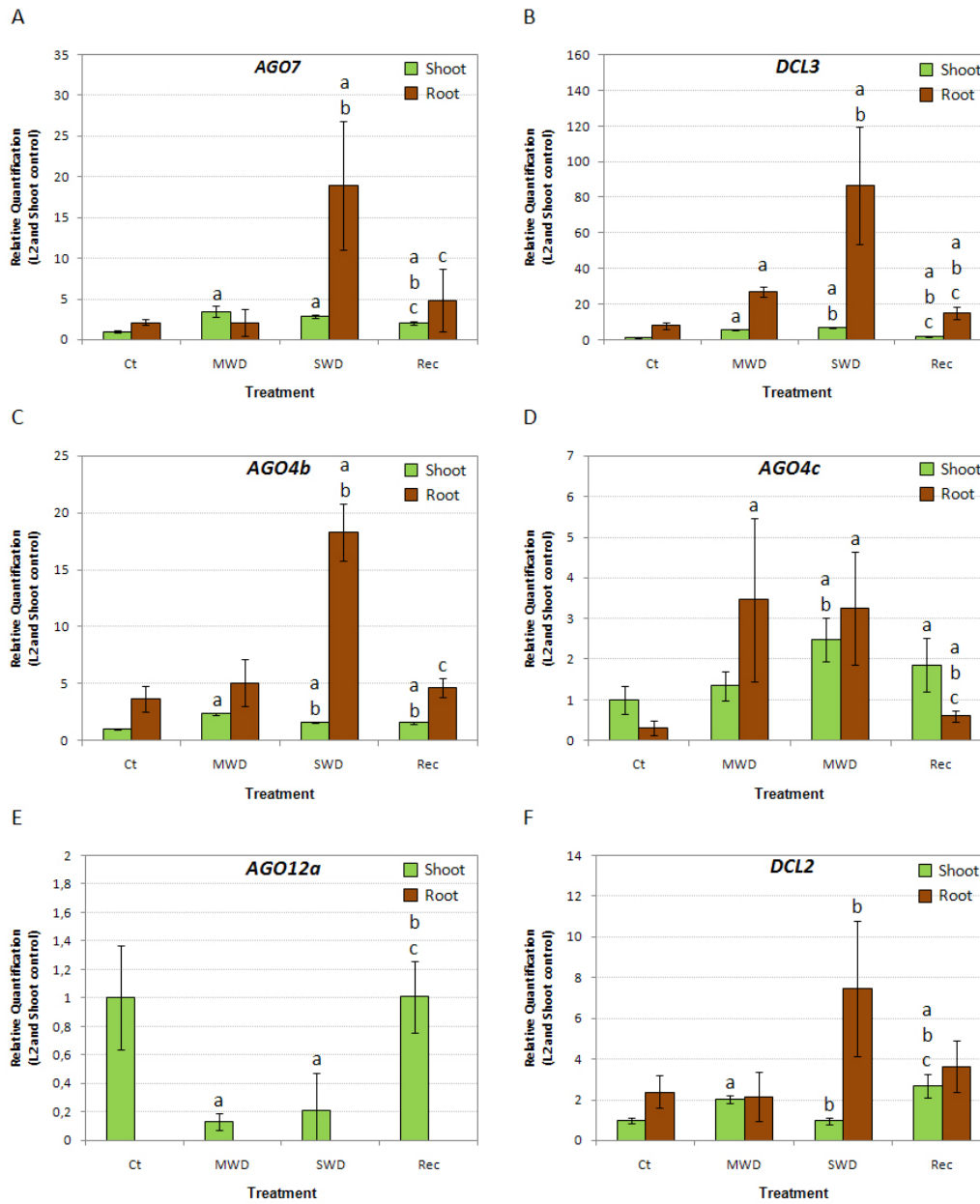


Figure 16: Relative expression levels of Argonautes and Dicer-like mRNAs in *M. truncatula*, by quantitative Real Time PCR. The shoot (green) and roots (brown) of *M. truncatula* are the tissues analyzed in the different water treatment conditions imposed to the plants. Values are the mean of two technical replicates of three independent cDNAs for each treatment (giving a total of 6 technical replicates) and bars represent standard errors. The relative expression was calculated using L2 as reference gene and normalized against to the shoot control treatment (Equation 5, Methods). The AGO1 and DCL1 primer pair was designed to give one amplicon with the cleavage site of their corresponding miRNA. A Two-tailed t-test of significance was applied, and was considered a significant difference with a p-value<0.05: a, significant difference with control of the same tissue, b, with MWD, and c, with SWD. (A) AGO7, Argonaute 7; (B) DCL3, Dicer-like 3; (C) AGO4b, Argonaute 4b; (D) AGO4c, Argonaute 4c; (E) AGO12a, Argonaute 12a; (F) DCL2, Dicer-like 2. Ct, Control; MWD, Moderate Water Deficit; SWD, Severe Water Deficit, Rec, Recovery.

To interrogate which of the MtAGOs included the conserved catalytic residues and could potentially act as the slicer component of the silencing effector complexes, we aligned the Piwi domains of all the MtAGOs and AtAGOs using CLUSTALW (Supplementary Figure 8). Two proteins, MtAGO1 and MtAGO7 were found to have the conserved domain DDH/H (Figure 14, C) in other cases like MtAGO12b the motif was missing or the residue H800 substituted by A, S or P, or in MtAGO2 the H (in the DDH motif) is substituted by one D (passing to a DDD motif), characteristic of AGO2 and AGO3 proteins in *A. thaliana* and *O. sativa* (Kapoor et al, 2008).

9.2.3. qRT-PCR of the AGOs and DCLs

DCL2 cleaves double-stranded virus RNA producing 22nt small RNAs (Vaucheret et al, 2008), and DCL3 cleaves endogenous dsRNA producing 24nt small RNAs involved in chromatin rearrangements (Vaucheret et al, 2008). AGO7 is involved in the biogenesis of trans-acting small RNAs (tas-siRNAs) derived from TAS3 mRNA (Vaucheret et al, 2008). AGO4 and AGO6 are involved in the transcriptional gene silencing (TGS) pathway, which perform chromatin remodeling (Vaucheret et al, 2008).

We performed the expression analysis by qRT-PCR of, MtAGO7, MtDCL3, MtAGO4b, MtAGO4c, MtAGO12a, MtDCL2 and MtAGO11 in shoots and roots of plants subjected to water deficit (Figure 16). MtAGO11 could not be quantified because the CT values were close to 40, and this is unacceptable for quantification (Bustin et al, 2009). MtAGO7, MtDCL3, MtAGO4b and MtAGO4c increased their expression in moderate and severe water deficit treatments in both tissues (Figure 16, A, B, C and D), the first three genes were more expressed in the roots than in the shoots and in MtAGO4c the expression levels were similar in shoots and roots. MtAGO12a expression was only found only in the shoots, decreasing its mRNA level in MWD and SWD (Figure 16, E). MtDCL2 expression increases just in SWD in the roots (Figure 16, F).

9.3. Discussion

9.3.1. Catalytic domain of AGOs

The slicing activity has been demonstrated for *A. thaliana* AGO1, AGO2, AGO4 and AGO7 (Baumberger and Baulcombe, 2005; Qi et al, 2005; Qi et al, 2006; Montgomery et al, 2008) and all of them, except AGO2, have an active site carrying a DDH motif (for AGO2 a DDD motif) (Figure 15, C) related with the Argonaute slicer function in plants and animals (Rivas et al, 2005). Baumberger and Baulcombe (2005) showed that the histidine residue in position 800 was essential for the cleavage activity of Argonaute1, but AtAGO4 has a slicing activity and have instead a serine residue. This means that the existence of this residue in Argonautes could not be so critical for their activity. MtAGO1, MtAGO2 and MtAGO7 have the motif DDH/H or DDD/H (in the case of AGO2) (Figure 15, C) and are homologous to AtAGO1, AtAGO2 and AtAGO7 indicating that they may have a slicer activity. In the cases of the MtAGO4a, b, c and d, AGO12a with a DDH/(A/S/P) and AGO11 with a DDD/S motif, the possibility of them to have a slicing activity can not be discarded. And in some organisms like in *Homo sapiens*, the AGO3 protein has a DDH domain but does not have slicing activity (Liu et al, 2004; Meister et al, 2004). On the other hand, the *Drosophila melanogaster* Piwi

domain has one DDK motif and has catalytic activity (Saito et al, 2006). This means that the existence or absence of a DDH domain does not preclude the protein from having or not slicing activity.

9.3.2. AGO7 and its involvement in water deficit response in *M. truncatula*

Argonaute 7 binds specifically with miR390 and directs the cleavage at 3' end of its non-coding target TAS3 RNA (Allen et al, 2005; Montgomery et al, 2008). The TAS3 cleavage products are stabilized by Suppressor of Gene Silencing 3 (SGS3), and one of the two TAS3 cleavage products is converted to dsRNA by RNA dependent RNA Polymerase 6 (RDR6) and finally diced by DCL4 into 21nts trans acting siRNA (ta-siRNA) a process assisted by a dsRNA binding protein 4 (DRB4) (Peragine et al, 2004; Vazquez et al, 2004; Gascioli et al, 2005; Yoshikawa et al, 2005; Adenot et al, 2006). The bioinformatic search for DCLs in *M. truncatula*, could not find one homolog sequence with *A. thaliana* DCL4, possibly because the *M. truncatula* genome is not yet fully sequenced. Three genes annotated as similar with Auxin Response factor 3 (ARF3) of *A. thaliana* were identified for *M. truncatula* as targets of two TAS3-derived ta-siRNAs (Jagadeeswaran et al, 2009).

A. thaliana ago7 mutants accelerate the juvenile to adult transition but not the onset of reproductive competence or flowering time (Hunter et al, 2003). ARF3 over-expression resulted in further acceleration of phase change and severe morphological and patterning defects of leaves and floral organs (Fahlgren et al, 2006). We observed that MtAGO7 is up-regulated in water deficit in both shoots and roots (Figure 16, A). Previous works correlated the role of TAS3 derived ta-siRNAs with plant development processes, but our results indirectly suggest that ARF3 can have a role in plant reaction to water deficit, maybe by repressing the development processes in the plant under stress conditions. The expression levels of the TAS3, the three ARF3 genes and TAS3 derived ta-siRNAs from *M. truncatula* should be analyzed under water deficit to test this hypothesis.

9.3.3. Water deficit response in *M. truncatula* and chromatin rearrangements

In plants, chromatin gene silencing involves the coordinated action of: (1) a DNA dependent RNA polymerase IV (Pol Iva and Pol IVb) that transcribes genomic regions with transposons and highly repeated sequences (Herr et al, 2005; Pontier et al, 2005); (2) a RNA dependent RNA polymerase 2 (RDR2), that produces a dsRNA sequence from the regions transcribed by RNA Pol IV (Chan et al, 2004); (3) a Dicer-like 3 (DCL3), that cleaves the dsRNA in 24nt siRNAs (Chan et al, 2004; Xie et al, 2004; Qi et al, 2005); (4) Argonaute 4 and 6 (AGO4 and AGO6), important for the accumulation of specific heterochromatin-related siRNAs, and for DNA methylation and subsequent transcriptional gene silencing (Zilberman et al, 2003; Chan et al, 2004; Zheng et al, 2007). In this work we showed that the MtDCL3, MtAGO4b and MtAGO4c are up-regulated in water deficit situations (Figure 16, B, C and D) which suggest that chromatin rearrangements may have a role in the plants response to water deficit.

9.3.4. MtAGO11 and MtAGO12a

MtAGO11 have similarities to AtAGO4 and AtAGO6 that are involved in DNA methylation and also with AtAGO8 and AtAGO9, but their functions remain to be discovered (Vaucheret et al, 2008). The

MtAGO11 sequence was initially annotated as being two proteins (CT030192_30 and CT030192_31) but using Fgenesh-H software we showed that it is just one gene and one protein (Supplementary Figure 4). MtAGO11 is expressed at very low levels in shoots and roots under all water deficit treatments with CTs values close to 40 that did not allow a good quantification of its expression. AtAGO8 was determined to be a pseudogene of AtAGO9, and has a very low expression level contrary to AtAGO9 that is highly transcribed (Vaucheret et al, 2008). It is possible that MtAGO11 is a pseudogene of MtAGO4 or may instead be expressed in other tissues not here analyzed (like MtAGO12a that is expressed in the shoots and not in the roots) or in other stress conditions imposed to the plant (Kapoor et al, 2008). To test this must be performed the expression analysis of MtAGO11 in other tissues (like flowers, seedpods, seedlings and nodules) and in plants under other stress conditions (like salt stress, heat stress or light stress).

MtAGO12a was just expressed in shoots, reduced its expression in water deficit conditions and returned to control levels in recovered plants (Figure 16, E). This differential expression in water deficit conditions suggests that MtAGO12a may have role in *M. truncatula* adaptation to water deficit. No information about the expression of MtAGO12 was found in the literature.

10. Conclusions

In this work we showed that the molecular responses to water deficit in *M. truncatula* were stronger in the roots comparing to the shoots, with the genes here analyzed. This is in conformity with the fact, that the roots are the plant organ that contact directly with the water deprivation in the soil and should perceive the stress stimulus at first.

With this work we were able to establish L2 as a good reference gene for studies involving water deficit responses in *M. truncatula*. We could show a relation between the RWC values (physiological parameter) of the plants in the 4 water status conditions and the expression level of water deficit-responsive genes ERD1, PIP1;5, Dehydrin and MI-1-PS (molecular/cellular parameter). The expression levels of ERD1, PIP1;5 and Dehydrin in shoots and roots and of MI-1-PS in roots were temporally and inversely related with the RWC values of the plants, confirming that the progressive water deficit induced an increase in transcription of drought-responsive genes.

In the second part of the work we show that *A. thaliana*, *P. trichocarpa* and *M. truncatula* share a common ancestor that had just one copy of the MIR168 gene. In *M. truncatula* genome there are 3 possible genes for MIR168 and that in its evolutionary lineage the first duplication event occurred between MIR168a and MIR168b/c, and the second between MIR168b and MIR168c.

We observed an inverse relation between the expression levels of the miR162 and its target Dicer-like 1 mRNA in *M. truncatula* roots under water deficit. From this we concluded that the down-regulation of miR162 in response to water deficit should decrease the number of DCL1 mRNA molecules cleaved, thus increasing the levels of its mRNA in the root cells.

It was not possible to relate the expression of miR168a/b and AGO1 in the different water status although both of them showed to be differentially expressed in *M. truncatula* under water deficit, principally in the roots. Then the role of miR168a/b in the plant adaptation to water deficit remains to be elucidated.

In last, we identified 3 Dicer-like and 9 Argonaute genes in the *M. truncatula* genome. The proteins DCL3, AGO4b, AGO4c (involved in chromatin rearrangements) and AGO7 (implicated in the production of ta-siRNAs) are being up-regulated in water deficit conditions showing that other silencing pathways besides the miRNA biosynthesis are regulated in *M. truncatula* under water deficit. These results highlight the importance of a better understanding of the small RNA silencing regulatory pathways in plants under water deprivation.

11. Future Perspectives

This work can be used in future studies:

1. It will be very interesting to verify if a linear correlation between the plant relative water content and the expression level of ERD1, PIP1;5 and Dehydrin genes in shoots and roots and MI-1-PS in roots can be established. For these studies biological replicates must be used in the quantification of both variables, and more points are needed for plant RWC.
2. We have shown that in roots under SWD the expression of MI-1-PS increase when compared to MWD, contrary to what happens in the shoots (Figure 7, D). A quantification of inositol and other inositol derived cyclitols levels in the roots would be a good start to comprehend why MI-1-PS was not inhibited in SWD. One hypothesis is that inositol could be transported in the xylem from the roots to the shoots, and then a quantification of its level in the xylem and phloem would be necessary to test this hypothesis.
3. We show that the expression of miR162 and miR168a/b are very similar in the different tissues (Figure 11) and in the water deficit conditions (Figure 12) except in the SWD in the roots comparing with the other conditions. A promoter analysis of MIR162, MIR168a and MIR168b should be performed to compare if they have a spacial and temporal relation of their expression in *M. truncatula*. The plants should be transformed with a construction of the β -glucuronidase (GUS) reporter gene expressed under the control of the promoter region of the MIR genes and the expression patterns analyzed and interpreted.
4. MtAGO7 increase its expression in *M. truncatula* under water deficit (Figure 16, A). We hypothesized that this would increase the number of ta-siRNAs derived from TAS3 mRNA. Finally it would increase the cleavage of their targets, the ARF3 mRNA, reducing their levels. The expression levels of the TAS3, the three ARF3 genes and TAS3 derived ta-siRNAs from *M. truncatula* should be analyzed under water deficit to test this hypothesis.
5. To try to understand the function of MtAGO12a, in situ hybridization of MtAGO12a mRNA in shoots could be done to know if its expression is tissue specific. Sequencing the small RNAs associated with MtAGO12a could also be done to try to understand its possible function (for example, AGO7 is associated predominantly with miR390, Montgomery et al, 2008). To achieve this MtAGO12a must be expressed with a fusion tag that allows its purification by immunoprecipitation together with the small RNAs associated with it.
6. Sequencing the small RNAs associated with the MtAGO4b and MtAGO4c to find if both genes are associated with the same RNAs or if each one is specific for different RNAs.

12. Bibliography

- Adai A, Johnson C, Mlotshwa S, Archer-Evans S, Manocha V, Vance V, Sundaresan V (2005) Computational prediction of miRNAs in *Arabidopsis thaliana*. *Genome Research*, 15:78–91.
- Adenot X, Elmayan T, Lauressergues D, Boutet S, Bouché N, Gascioli V, Vaucheret H (2006) DRB4-dependent TAS3 trans-acting siRNAs control leaf morphology through AGO7. *Current Biology*, 16(9):927-932.
- Alexandersson E, Fraysse L, Sjövall-Larsen S, Gustavsson S, Fellert M, Karlsson M, Johanson U, Kjellbom P (2005) Whole gene family expression and drought stress regulation of aquaporins. *Plant Molecular Biology* 59(3):469-484.
- Allagulova ChR, Gimalov FR, Shakirova FM, Vakhitov VA (2003) The plant dehydrins: structure and putative functions. *Biochemistry (Moscow)*, 68(9):945-951.
- Allen E, Xie Z, Gustafson AM, Sung GH, Spatafora JW, Carrington JC (2004) Evolution of microRNA genes by inverted duplication of target gene sequences in *Arabidopsis thaliana*. *Nature Genetics*, 36(12):1282-1290.
- Allen E, Xie Z, Gustafson AM, Carrington JC (2005) microRNA-directed phasing during trans-acting siRNA biogenesis in plants. *Cell*, 121(2):207-221.
- Ambros V, Bartel B, Bartel DP, Burge CB, Carrington JC, Chen X, Dreyfuss G, Eddy SR, Griffiths-Jones S, Marshall M, Matzke M, Ruvkun G, Tuschl T (2008) A uniform system for microRNA annotation. *RNA*, 9:277-279.
- Andersen CL, Jensen JL, Orntoft TF (2004) Normalisation of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalisation, applied to bladder and colon cancer data sets. *Cancer Research*, 64:5245–525
- Araújo SS, Duque ASRLA, Santos DMMF, Feveiro MPS (2004) An efficient transformation method to regenerate a high number of transgenic plants using a new embryogenic line of *Medicago truncatula* cv. Jemalong. *Plant Cell, Tissue and Organ Culture*, 78(2): 123-131.
- Aukerman MJ, Sakai H (2003) Regulation of flowering time and floral organ identity by a MicroRNA and its APETALA2-like target genes. *The Plant Cell*, 15:2730–2741.
- Barker DG, Bianchi S, Blondon F, Datteé Y, Duc G et al. (1990) *Medicago truncatula*, a model plant for studying the molecular genetics of the Rhizobium-legume symbiosis. *Plant Molecular Biology Reporter*, 8:40-49.
- Bartel B, Bartel DP (2003) MicroRNAs: at the root of plant development? *Plant Physiology*, 132(2):709-717.
- Baumberger N, Baulcombe DC (2005) *Arabidopsis* ARGONAUTE1 is an RNA Slicer that selectively recruits microRNAs and short interfering RNAs. *The Proceedings of the National Academy of Sciences*, 102(33):11928-11933.
- Bennett MD, Leitch IJ (1995) Nuclear DNA amounts in angiosperms. *Annals of Botany*, 76:113-176.
- Bonnet E, Wuyts J, Rouze P, Van de Peer Y (2004) Detection of 91 potential conserved plant microRNAs in *Arabidopsis thaliana* and *Oryza sativa* identifies important target genes. *Proceedings of the National Academy of Sciences USA*, 101:11511–11516.
- Bota J, Flexas J, Medrano H (2004) Is photosynthesis limited by decreased Rubisco activity and RuBP content under progressive water stress? *New Phytologist*, 162:671–681.
- Bray EA (1993) Molecular Responses to Water Deficit. *Plant Physiology*, 103(4):1035-1040.
- Brodersen P, Sakvarelidze-Achard L, Bruun-Rasmussen M, Dunoyer P, Yamamoto YY, Sieburth L, Voinnet O (2008) Widespread translational inhibition by plant miRNAs and siRNAs. *Science*, 320:1185–1190.
- Bustin SA, Benes V, Garson JA, Hellems J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, Wittwer CT (2009) The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clinical Chemistry*, 55(4):611-622.
- Carels N, Bernardi G (2000) Two classes of genes in plants. *Genetics*, 154(4):1819-1825.
- Čatský, J (1960) Determination of water deficit in disks cut out from leaf blades. *Brief Communication*, 2(1):76-78.
- Cerutti L, Mian N, Bateman A (2000) Domains in gene silencing and cell differentiation proteins: the novel PAZ domain and redefinition of the Piwi domain. *Trends in Biochemical Sciences*, 25(10):481-482.
- Chan SW, Zilberman D, Xie Z, Johansen LK, Carrington JC, Jacobsen SE (2004) RNA silencing genes control de novo DNA methylation. *Science*, 303(5662):1336.
- Chaves MM (1991) Effects of water deficit on carbon assimilation. *Journal of Experimental Botany*, 42:1–16..
- Chaves MM, Maroco JP, Pereira JS (2003) Understanding plant responses to drought — from genes to the whole plant. *Functional Plant Biology*, 30(3):239-264.

- Chen X, Liu J, Cheng Y, Jia D (2002) HEN1 functions pleiotropically in Arabidopsis development and acts in C function in the flower. *Development*, 129:1085–1094.
- Chen X (2004) A microRNA as a translational repressor of APETALA2 in Arabidopsis flower development. *Science*, 303:2022–2025.
- Chen X (2008) MicroRNA metabolism in plants. *Current Topics in Microbiology and Immunology*, 320:117-136.
- Choi H, Hong J, Ha J, Kang J, Kim SY (2000) ABFs, a family of ABA-responsive element binding factors. *The Journal of Biological Chemistry*, 275(3):1723-1730.
- Cruz de Carvalho MH, Laffray D, Louguet P (1998) Comparison of the physiological responses of *Phaseolus vulgaris* and *Vigna unguiculata* cultivars when submitted to drought conditions. *Environmental and experimental botany*, 40:197–207.
- Dezulian T, Palatnik JF, Huson D, Weigel D (2005) Conservation and divergence of microRNA families in plants. *Genome Biology*, 6:P13.
- Diakić M (2006) DUF283 domain of Dicer proteins has a double-stranded RNA-binding fold. *Bioinformatics*, 22(22):2711-2714.
- Dure L, Crouch M, Harada J, David T, Mundy J, Quatrano R, Thomas T, Sung ZR (1989) Common amino acid sequence domains among the LEA proteins of higher plants. *Plant Molecular Biology*, 12(5):475-486.
- Eamens A, Wang MB, Smith NA, Waterhouse PM (2008) NA silencing in plants: yesterday, today, and tomorrow. *Plant Physiology*, 147(2):456-468.
- Fahlgren N, Montgomery TA, Howell MD, Allen E, Dvorak SK, Alexander AL, Carrington JC (2006) Regulation of AUXIN RESPONSE FACTOR3 by TAS3 ta-siRNA affects developmental timing and patterning in Arabidopsis. *Current Biology*, 16(9):939-944.
- Fang Y, Spector, DL (2007) Identification of nuclear dicing bodies containing proteins for microRNA biogenesis in living Arabidopsis plants. *Current Biology*, 17:818–823.
- Fujii H, Chiou TJ, Lin SI, Aung K, Zhu JK (2005) A miRNA involved in phosphate-starvation response in Arabidopsis. *Current Biology*, 15(22):2038-2043.
- Garcia D (2008) A miRacle in plant development: role of microRNAs in cell differentiation and patterning. *Seminars in Cell and Developmental Biology*, 19(6):586-595.
- Gascioli 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. *Current Biology*, 15(16):1494-1500.
- Gazzani S, Li M, Maistri S, Scarponi E, Graziola M, Barbaro E, Wunder J, Furini A, Saedler H, Varotto C (2009) Evolution of MIR168 paralogs in Brassicaceae. *BMC Evolutionary Biology*, 9:62.
- Griffiths-Jones S, Grocock RJ, van Dongen S, Bateman A, Enright AJ (2006) miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Research*, 34(Database Issue):D140-D144.
- Halperin T, Adam Z (1996) Degradation of mistargeted OEE33 in the chloroplast stroma. *Plant Molecular Biology*, 30(5):925-933.
- Hellwege EM, Dietz K, Volk OH, Hartung W (2004) Abscisic acid and the induction of desiccation tolerance in the extremely xerophilic liverwort *Exormothesa holstii*. *Planta*, 194(4):525-531.
- Herr AJ, Jensen MB, Dalmay T, Baulcombe DC (2005) RNA polymerase IV directs silencing of endogenous DNA. *Science*, 308(5718):118-120.
- Hofacker IL (2003) Vienna RNA secondary structure server. *Nucleic Acids Research*, 31(13): 3429–3431.
- Hong-Bo S, Zong-Suo L, Ming-An S (2005) LEA proteins in higher plants: structure, function, gene expression and regulation. *Colloids and Surfaces B: Biointerfaces*, 45(3-4):131-5.
- Hsiao TC (1973) Plant Responses to Water Stress. *Annual Review of Plant Physiology*, 24:519-570.
- Huang S, An YQ, McDowell JM, McKinney EC, Meagher RB (1997) The Arabidopsis ACT11 actin gene is strongly expressed in tissues of the emerging inflorescence, pollen, and developing ovules. *Plant Molecular Biology*, 33(1):125-139.
- Hunter C, Sun H, Poethig RS (2003) The Arabidopsis heterochronic gene ZIPPY is an ARGONAUTE family member. *Current Biology*, 13(19):1734-1739.
- Hutvagner G, Simard MJ (2008) Argonaute proteins: key players in RNA silencing. *Nature Reviews Molecular Cell Biology*, 9:22–32.
- Ingram J, Bartels D (1996) THE MOLECULAR BASIS OF DEHYDRATION TOLERANCE IN PLANTS. *Annual Review of Plant Physiology and Plant Molecular Biology*, 47:377-403.

- Jagadeeswaran G, Zheng Y, Li YF, Shukla LI, Matts J, Hoyt P, Macmil SL, Wiley GB, Roe BA, Zhang W, Sunkar R (2009) Cloning and characterization of small RNAs from *Medicago truncatula* reveals four novel legume-specific microRNA families. *New Phytologist*.
- Jinek M, Doudna JA (2009) A three-dimensional view of the molecular machinery of RNA interference. *Nature*, 457(7228):405-412.
- Johanson U, Karlsson M, Johansson I, Gustavsson S, Sjövall S, Fraysse L, Weig AR, Kjellbom P (2001) The complete set of genes encoding major intrinsic proteins in *Arabidopsis* provides a framework for a new nomenclature for major intrinsic proteins in plants. *Plant Physiology*, 126:1358–69.
- Johnson MD, Sussex IM (1995) 1 L-myo-Inositol 1-Phosphate Synthase from *Arabidopsis thaliana*. *Plant Physiology*, 107(2):613–619.
- Jones-Rhoades MW, Bartel DP (2004) Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. *Molecular Cell*, 14(6):787-799.
- Juarez MT, Kui JS, Thomas J, Heller BA, Timmermans MC (2004) microRNA-mediated repression of *rolled leaf1* specifies maize leaf polarity. *Nature*, 428:84–88.
- Kandasamy MK, Gilliland LU, McKinney EC, Meagher RB (2001) One plant actin isoform, ACT7, is induced by auxin and required for normal callus formation. *The Plant Cell*, 13(7):1541-1554.
- Kapoor M, Arora R, Lama T, Nijhawan A, Khurana JP, Tyagi AK, Kapoor S (2008) Genome-wide identification, organization and phylogenetic analysis of Dicer-like, Argonaute and RNA-dependent RNA Polymerase gene families and their expression analysis during reproductive development and stress in rice. *BMC Genomics*, 9:451.
- Khvorova A, Reynolds A, Jayasena SD (2003) Functional siRNAs and miRNAs exhibit strand bias. *Cell*, 115:209–216.
- Kidner CA, Martienssen RA (2004) Spatially restricted microRNA directs leaf polarity through ARGONAUTE1. *Nature*, 428:81–84.
- Kim S, Hwang do W, Lee DS (2009) A study of microRNAs in silico and in vivo: bioimaging of microRNA biogenesis and regulation. *FEBS Journal*, 276(8):2165-74.
- Kiyosue T, Yamaguchi-Shinozaki K, Shinozaki K (1993) Characterization of cDNA for a dehydration-inducible gene that encodes a CLP A, B-like protein in *Arabidopsis thaliana* L. *Biochemical and Biophysical Research Communications*, 196(3):1214-1220.
- Kurihara Y, Takashi Y, Watanabe Y (2006) The interaction between DCL1 and HYL1 is important for efficient and precise processing of pri-miRNA in plant microRNA biogenesis. *RNA*, 12:206–212.
- Kurihara Y, Watanabe Y (2004) *Arabidopsis* micro-RNA biogenesis through Dicer-like 1 protein functions. *Proceedings of the National Academy of Sciences USA*, 101:12753–12758.
- Lanet E, Delannoy E, Sormani R, Floris M, Brodersen P, Crété P, Voinnet O, Robaglia C (2009) Biochemical evidence for translational repression by *Arabidopsis* microRNAs. *Plant Cell*, 21(6):1624.
- Lawlor DW, Cornic G (2002) Photosynthetic carbon assimilation and associated metabolism in relation to water deficits in higher plants. *Plant, Cell & Environment*, 25:275–294.
- 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(5):843-854.
- Lee Y, Kim M, Han J, Yeom KH, Lee S, Baek SH, Kim VN (2004) MicroRNA genes are transcribed by RNA polymerase II. *The EMBO Journal*, 23(20):4051-4060.
- Li J, Yang Z, Yu B, Liu J, Chen X (2005) Methylation protects miRNAs and siRNAs from a 3' end uridylation activity in *Arabidopsis*. *Current Biology*, 15:1501–1507.
- Liu B, Li P, Li X, Liu C, Cao S, Chu C, Cao X (2005) Loss of function of *OsDCL1* affects microRNA accumulation and causes developmental defects in rice. *Plant Physiology*, 139:296–305.
- Liu J, Carmell MA, Rivas FV, Marsden CG, Thomson JM, Song JJ, Hammond SM, Joshua-Tor L, Hannon GJ (2004) *Argonaute2* is the catalytic engine of mammalian RNAi. *Science*, 305(5689):1437-1441.
- Liu Q, Zhang YC, Wang CY, Luo YC, Huang QJ, Chen SY, Zhou H, Qu LH, Chen YQ (2009a) Expression analysis of phytohormone-regulated microRNAs in rice, implying their regulation roles in plant hormone signaling. *FEBS Letters*, 583(4):723-728.
- Liu Q, Feng Y, Zhu Z (2009b) Dicer-like (DCL) proteins in plants. *Functional & Integrative Genomics*, 9(3):277-286.
- Llave C, Xie Z, Kasschau KD, Carrington JC (2002) Cleavage of Scarecrow-like mRNA targets directed by a class of *Arabidopsis* miRNA. *Science*, 297:2053–2056.

- Loewus FA, Murthy PPN (2000) myo-Inositol metabolism in plants. *Plant Science*, 150(1):1-19.
- Lu S, Sun YH, Shi R, Clark C, Li L, Chiang VL (2005) Novel and mechanical stress-responsive MicroRNAs in *Populus trichocarpa* that are absent from *Arabidopsis*. *The Plant Cell*, 17(8):2186-2203.
- Mahajan S, Tuteja N (2005) Cold, salinity and drought stresses: an overview. *Archives of Biochemistry and Biophysics* 444(2):139-158.
- Majeran F, Wollman F, Vallon O (2000) Evidence for a Role of ClpP in the Degradation of the Chloroplast Cytochrome b6f Complex. *The Plant Cell* 12(1):137–150.
- Majumder AL, Johnson MD, Henry SA (1997) 1L-myo-inositol-1-phosphate synthase. *Biochimica et Biophysica Acta*, 1348(1-2):245–256.
- Marchler-Bauer A, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, Fong JH, Geer LY, Geer RC, Gonzales NR, Gwadz M, He S, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Liebert CA, Liu C, Lu F, Lu S, Marchler GH, Mullokandov M, Song JS, Tasneem A, Thanki N, Yamashita RA, Zhang D, Zhang N, Bryant SH (2009) CDD: specific functional annotation with the Conserved Domain Database. *Nucleic Acids Research*, 37(Database issue):D205-10.
- Margis R, Fusaro AF, Smith NA, Curtin SJ, Watson JM, Finnegan EJ, Waterhouse PM (2006) The evolution and diversification of Dicers in plants. *FEBS Letters*, 580(10):2442-2450.
- Marty I, Meyer Y (1992) cDNA nucleotide sequence and expression of a tobacco cytoplasmic ribosomal protein L2 gene. *Nucleic Acids Research*, 20(7):1517–1522.
- Matranga C, Tomari Y, Shin C, Bartel DP, Zamore PD (2005) Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes. *Cell*, 123:607–620.
- Maurel C, Verdoucq L, Luu DT, Santoni V (2008) Plant aquaporins: membrane channels with multiple integrated functions. *Annual Review of Plant Biology*, 59:595-624.
- Megraw M, Baev V, Rusinov V, Jensen ST, Kalantidis K, Hatzigeorgiou AG (2006) MicroRNA promoter element discovery in *Arabidopsis*. *RNA*, 12(9):1612-1619.
- Meister G, Landthaler M, Patkaniowska A, Dorsett Y, Teng G, Tuschl T (2004) Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Molecular Cell*, 15(2):185-197.
- Mi S, Cai T, Hu Y, Chen Y, Hodges E, Ni F, Wu L, Li S, Zhou H, Long C, Chen S, Hannon GJ, Qi Y (2008) Sorting of small RNAs into *Arabidopsis* argonaute complexes is directed by the 5' terminal nucleotide. *Cell*, 133(1):116-127.
- Millar AA, Waterhouse PM (2005) Plant and animal microRNAs: similarities and differences. *Functional & Integrative Genomics*, 5(3):129-135.
- Miyoshi K, Tsukumo H, Nagami T, Siomi H, Siomi MC (2005) Slicer function of *Drosophila* Argonautes and its involvement in RISC formation. *Genes & Development*, 19:2837–2848.
- Moffatt BA, Pthe C, Laloue M (1991) Metabolism of benzyladenine is impaired in mutant of *Arabidopsis thaliana* lacking adenine phosphoribosyltransferase activity. *Plant Physiology*, 95:900–908.
- Montgomery TA, Howell MD, Cuperus JT, Li D, Hansen JE, Alexander AL, Chapman EJ, Fahlgren N, Allen E, Carrington JC (2008) Specificity of ARGONAUTE7-miR390 interaction and dual functionality in TAS3 trans-acting siRNA formation. *Cell*, 133(1):128-141.
- Moxon S, Schwach F, MacLean D, Dalmay T, Studholme DJ, Moulton V (2008) A toolkit for analysing large-scale plant small RNA datasets. *Bioinformatics*, 24(19):2252-2253.
- Nakashima K, Kiyosue T, Yamaguchi-Shinozaki K, Shinozaki K (1997) A nuclear gene *erd1* encoding a chloroplast-targeted Clp protease regulatory subunit homolog is not only induced by water stress but also developmentally up-regulated during senescence in *Arabidopsis thaliana*. *Plant Journal* 12(4):851-861.
- Navarro L, Dunoyer P, Jay F, Arnold B, Dharmasiri N, Estelle M, Voinnet O, Jones JD (2006) A plant miRNA contributes to antibacterial resistance by repressing auxin signaling. *Science*, 312(5772):436-439.
- Nelson DE, Koukoumanos M, Bohnert HJ (1999) Myo-inositol-dependent sodium uptake in ice plant. *Plant Physiology*, 119(1):165-172.
- Nicot N, Hausman JF, Hoffmann L, Evers D (2005) Housekeeping gene selection for real-time RT-PCR normalization in potato during biotic and abiotic stress. *Journal of Experimental Botany*, 56(421):2907-2914.
- Notredame C, Higgins DG, Heringa J (2000) T-Coffee: A novel method for multiple sequence alignments. *Journal of Molecular Biology*, 302(1):205-17.
- Nunes C, Araújo SS, Silva JM, Fevereiro MPS, Silva AB (2008) Physiological responses of the legume model *Medicago truncatula* cv. Jemalong to water deficit. *Environmental and Experimental Botany*, 63(1-3):289-296.

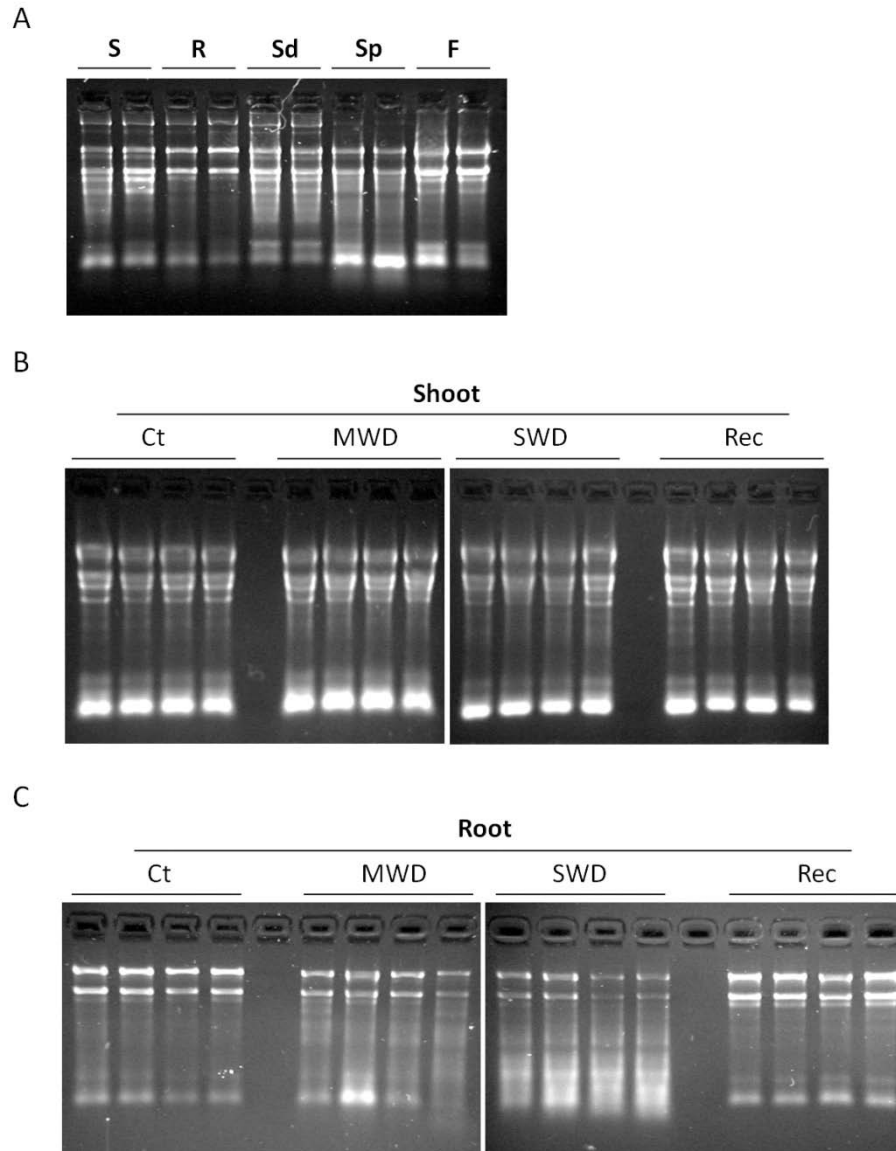
- Olmedo G, Guzmán P (2008) Processing precursors with RNase III in plants. *Plant Science*, 175(6):741-746.
- Pall GS, Codony-Servat C, Byrne J, Ritchie L, Hamilton A (2007) Carbodiimide-mediated cross-linking of RNA to nylon membranes improves the detection of siRNA, miRNA and piRNA by northern blot. *Nucleic Acids Research*, 35(8):e60.
- Park MY, Wu G, Gonzalez-Sulser A, Vaucheret H, Poethig RS (2005) Nuclear processing and export of microRNAs in *Arabidopsis*. *Proceedings of the National Academy of Sciences USA*, 102:3691–3696.
- Peragine A, Yoshikawa M, Wu G, Albrecht HL, Poethig RS (2004) SGS3 and SGS2/SDE1/RDR6 are required for juvenile development and the production of trans-acting siRNAs in *Arabidopsis*. *Genes & Development* 18(19):2368-2379.
- Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Research*, 45(9): e45.
- Pontier D, Yahubyan G, Vega D, Bulski A, Saez-Vasquez J, Hakimi MA, Lerbs-Mache S, Colot V, Lagrange T (2005) Reinforcement of silencing at transposons and highly repeated sequences requires the concerted action of two distinct RNA polymerases IV in *Arabidopsis*. *Genes & Development* 19(17):2030-2040.
- Qi Y, Denli AM, Hannon GJ (2005) Biochemical specialization within *Arabidopsis* RNA silencing pathways (2005) Biochemical specialization within *Arabidopsis* RNA silencing pathways. *Molecular Cell*, 19(3):421-428.
- Qi Y, He X, Wang XJ, Kohany O, Jurka J, Hannon GJ (2006) Distinct catalytic and non-catalytic roles of ARGONAUTE4 in RNA-directed DNA methylation. *Nature*, 443(7114):1008-1012.
- Rajagopalan R, Vaucheret H, Trejo J, Bartel DP (2006) A diverse and evolutionarily fluid set of microRNAs in *Arabidopsis thaliana*. *Genes & Development*, 20(24):3407-3425.
- Ramachandran V, Chen X (2008) Degradation of microRNAs by a family of exoribonucleases in *Arabidopsis*. *Science*, 321:1490–1492.
- Reinhart BJ, Weinstein EG, Rhoades MW, Bartel B, Bartel DP (2002) MicroRNAs in plants. *Genes & Development*, 16:1616–1626.
- Rhoades MW, Reinhart BJ, Lim LP, Burge CB, Bartel B, Bartel DP (2002) Prediction of plant microRNA targets. *Cell*, 110:513–520.
- 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. *Nature Structural & Molecular Biology*, 12(4):340-349.
- Rose RJ (2008) *Medicago truncatula* as a model for understanding plant interactions with other organisms, plant development and stress biology : past, present and future. *Functional Plant Biology*, 35:253-264;
- Saito K, Nishida KM, Mori T, Kawamura Y, Miyoshi K, Nagami T, Siomi H, Siomi MC (2006) Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the *Drosophila* genome. *Genes & Development*, 20(16):2214-2222.
- Schwarz DS, Hutvagner G, Du T, Xu Z, Aronin N, Zamore PD (2003) Asymmetry in the assembly of the RNAi enzyme complex. *Cell*, 115:199–208.
- Scotti Campos P, Ramalho JC, Lauriano JA, Silva MJ, Céu Matos M (1999) Effects of drought on photosynthetic performance and water relations of four *Vigna* genotypes. *Photosynthetica*, 36:79–87.
- Seki M, Umezawa T, Urano K, Shinozaki K (2007) Regulatory metabolic networks in drought stress responses. *Current Opinion in Plant Biology*, 10(3):296-302.
- Shao HB, Chu LY, Shao MA, Zhao CX (2008) Advances in functional regulation mechanisms of plant aquaporins: their diversity, gene expression, localization, structure and roles in plant soil-water relations. *Molecular Membrane Biology*, 25(3):179-91.
- Sheveleva E, Chmara W, Bohnert HJ, Jensen RG (1997) Increased Salt and Drought Tolerance by D-Ononitol Production in Transgenic *Nicotiana tabacum* L. *Plant Physiology*, 115(3): 1211–1219.
- Shinozaki K, Yamaguchi-Shinozaki K (2007) Gene networks involved in drought stress response and tolerance. *Journal of Experimental Botany*, 58(2):221-227.
- Song JJ, Smith SK, Hannon GJ, Joshua-Tor L (2004) Crystal structure of Argonaute and its implications for RISC slicer activity, 305(5689):1434-1437.
- Sun Y, Carneiro N, Clore AM, Moro GL, Habben JE, Larkins BA (1997) Characterization of maize elongation factor 1A and its relationship to protein quality in the endosperm. *Plant Physiology*, 115(3):1101-1107.
- Sunkar R, Zhu JK (2004) Novel and stress-regulated microRNAs and other small RNAs from *Arabidopsis*. *The Plant Cell*, 16(8):2001-2019.
- Sunkar R, Kapoor A, Zhu JK (2006) Posttranscriptional induction of two Cu/Zn superoxide dismutase genes in *Arabidopsis* is mediated by downregulation of miR398 and important for oxidative stress tolerance. *Plant Cell*, 18:2051–2065.

- Szittya G, Moxon S, Santos DM, Jing R, Fevereiro MPS, Moulton V, Dalmay T (2008) High-throughput sequencing of *Medicago truncatula* short RNAs identifies eight new miRNA families. *BMC Genomics*, 9:593.
- Taiz L, Zeiger E (2002) *Plant Physiology*, 3rd edn. Sinauer Associates, Inc.
- Tam W (2001) Identification and characterization of human BIC, a gene on Chromosome 21 that encodes a noncoding RNA. *Gene*, 274:157–167.
- Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, 24:1596-1599.
- Tang G, Reinhart BJ, Bartel DP, Zamore PD (2003) A biochemical framework for RNA silencing in plants. *Genes & Development*, 17:49–63.
- Tanguilig VC, Yambao EB, O'toole JC, De Datta SK (1987) Water stress effects on leaf elongation, leaf water potential, transpiration, and nutrient uptake of rice, maize, and soybean. *Plant and Soil*, 103(2):155-168.
- Thompson JD, Higgins DG, Gibson TJ (1994) ClustalW—improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, 22:4673–4680.
- Tolia NH, Joshua-Tor L (2007) Slicer and the argonauts. *Nature Chemical Biology*, 3:36–43.
- Ueno Y, Ishikawa T, Watanabe K, Terakura S, Iwakawa H, Okada K, Machida C, Machida Y (2007) Histone deacetylases and ASYMMETRIC LEAVES2 are involved in the establishment of polarity in leaves of *Arabidopsis*. *The Plant Cell*, 19(2):445-457.
- Uno Y, Furihata T, Abe H, Yoshida R, Shinozaki K, Yamaguchi-Shinozaki K (2000) *Arabidopsis* basic leucine zipper transcription factors involved in an abscisic acid-dependent signal transduction pathway under drought and high-salinity conditions. *The Proceedings of the National Academy of Sciences*, 97(21):11632-11637.
- van Kan JAL, Joosten MHAJ, Wagemakers CAM, van den Berg-Velthuis GCM and Wit PJGM (1992) Differential accumulation of mRNAs encoding extracellular and intracellular PR proteins in tomato induced by virulent and avirulent races of *Cladosporium fulvum*. *Plant Molecular Biology*, 20: 513-527.
- Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F (2002) Accurate normalization of real-time geometric averaging of multiple internal control genes. *Genome Biol.* 3:research0034.1
- Várallyay E, Burgýán J, Havelda Z (2007) Detection of microRNAs by Northern blot analyses using LNA probes. *Methods*, 43:140-145.
- Vaucheret H, Vazquez F, Crété P, Bartel DP (2004) The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. *Genes & Development*, 18(10):1187-1197.
- Vaucheret H, Mallory AC, Bartel DP (2006) AGO1 homeostasis entails coexpression of MIR168 and AGO1 and preferential stabilization of miR168 by AGO1. *Molecular Cell*, 22(1):129-136.
- Vaucheret H (2008) Plant ARGONAUTES. *Trends in Plant Science*, 13(7):350-358.
- Vazquez F, Vaucheret H, Rajagopalan R, Lepers C, Gascioli V, Mallory AC, Hilbert JL, Bartel DP, Crété P (2004) Endogenous trans-acting siRNAs regulate the accumulation of *Arabidopsis* mRNAs. *Molecular Cell*, 16(1):69-79.
- Vazquez F, Blevins T, Ailhas J, Boller T, Meins F Jr (2008) Evolution of *Arabidopsis* MIR genes generates novel microRNA classes. *Nucleic Acids Research* 36(20):6429-6438.
- Voinnet O (2009) Origin, biogenesis, and activity of plant microRNAs. *Cell*, 136(4):669-687.
- Wang JW, Schwab R, Czech B, Mica E, Weigel D (2008) Dual effects of miR156-targeted SPL genes and CYP78A5/KLUH on plastochron length and organ size in *Arabidopsis thaliana*. *Plant Cell*, 20:1231–1243.
- Weaver LM, Froehlich JE, Amasino RM (1999) Chloroplast-targeted ERD1 protein declines but its mRNA increases during senescence in *Arabidopsis*. *Plant Physiology*, 119(4):1209-1216.
- Wu F, Yu L, Cao W, Mao Y, Liu Z, He Y (2007) The N-terminal double-stranded RNA binding domains of *Arabidopsis* HYPONASTIC LEAVES1 are sufficient for pre-microRNA processing. *The Plant Cell*, 19:914–925.
- Xie Z, Kasschau KD, Carrington JC (2003) Negative feedback regulation of Dicer-Like1 in *Arabidopsis* by microRNA-guided mRNA degradation. *Current Biology*, 13(9):784-789.
- 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 Biology*, 2(5):E104.
- Xie Z, Allen E, Fahlgren N, Calamar A, Givan SA, Carrington JC (2005) Expression of *Arabidopsis* MIRNA genes. *Plant Physiology*, 138(4):2145-2154.

- Yamaguchi-Shinozaki K, Shinozaki K (2006) Transcriptional regulatory networks in cellular responses and tolerance to dehydration and cold stresses. *Annual Review of Plant Biology*, 57:781-803.
- Yang Z, Ebright YW, Yu B, Chen X (2006) HEN1 recognizes 21–24 nt small RNA duplexes and deposits a methyl group onto the 2' OH of the 3' terminal nucleotide. *Nucleic Acids Research*, 34:667–675.
- Yoshikawa M, Peragine A, Park MY, Poethig RS (2005) A pathway for the biogenesis of trans-acting siRNAs in Arabidopsis. *Genes & Development*, 19(18):2164-2175.
- Young ND, Mudge J, Ellis TH (2003) Legume genomes: more than peas in a pod. *Current Opinion in Plant Biology*, 6(2):199-204.
- Young ND, Cannon SB, Sato S, Kim D, Cook DR, Town CD, Roe BA, Tabata S (2005) Sequencing the genespaces of *Medicago truncatula* and *Lotus japonicus*. *Plant Physiology*, 137:1174-1181.
- Yu B, Yang Z, Li J, Minakhina S, Yang M, Padgett RW, Steward R, Chen X (2005) Methylation as a crucial step in plant microRNA biogenesis. *Science*, 307:932–935.
- Zelazny E, Borst JW, Muylaert M, Batoko H, Hemminga MA, Chaumont F (2007) FRET imaging in living maize cells reveals that plasma membrane aquaporins interact to regulate their subcellular localization. *Proceedings of the National Academy of Sciences*, 104(30):12359-12364.
- Zhang B, Pan X, Anderson TA (2006a) Identification of 188 conserved maize microRNAs and their targets. *FEBS Letters*, 580(15):3753-3762.
- Zhang BH, Pan XP, Cox SB, Cobb GP, Anderson TA (2006b) Evidence that miRNAs are different from other RNAs. *Cellular and Molecular Life Sciences*, 63:246–254.
- Zhao B, Liang R, Ge L, Li W, Xiao H, Lin H, Ruan K, Jin Y (2007) Identification of drought-induced microRNAs in rice. *Biochemical and Biophysical Research Communications*, 354(2):585-590.
- Zhao S, Fernald RD (2005) Comprehensive algorithm for quantitative real-time polymerase chain reaction. *Journal of Computational Biology*, 12(8):1047-1064.
- Zheng X, Zhu J, Kapoor A, Zhu JK (2007) Role of Arabidopsis AGO6 in siRNA accumulation, DNA methylation and transcriptional gene silencing. *The EMBO Journal*, 26(6):1691-1701.
- Zhou X, Wang G, Zhang W (2007) UV-B responsive microRNA genes in Arabidopsis thaliana. *Molecular Systems Biology*, 3:103.
- Zilberman D, Cao X, Jacobsen SE (2003) ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. *Science*, 299(5607):716-719.

13. Supplementary Tables and Figures

13.1. Methods - Supplementary Tables and Figures



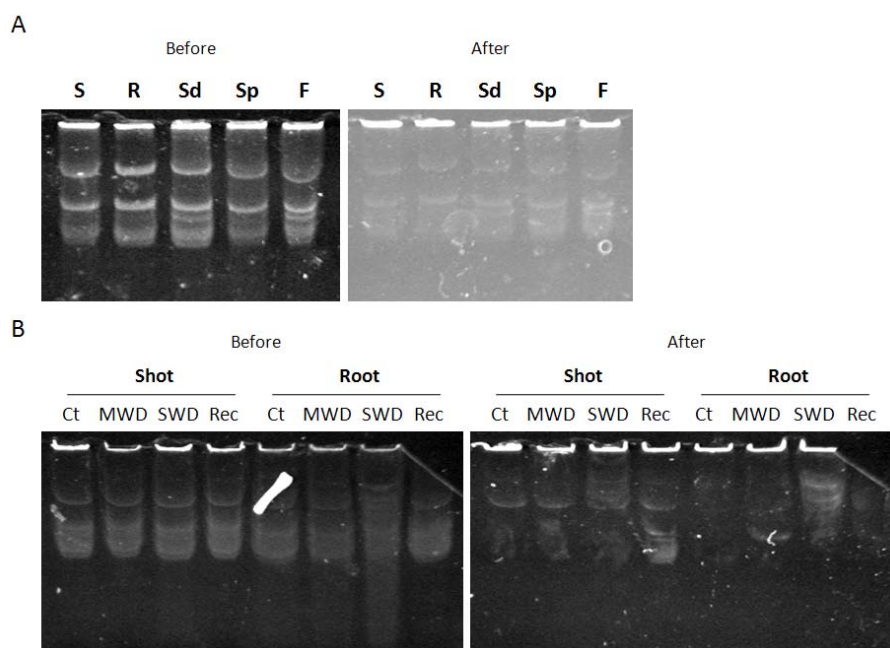
Supplementary Figure 1: Integrity of the RNA extracted from *M. truncatula*. To test the quality of the extracted RNA, it was run in a 2% agarose gel (0.5X TBE) stained with Sybr Safe. In this way is possible to see if exists degraded RNA. (A) RNA from shoots (S) and roots (R), seedlings (Sd), seed pods (Sp) and flowers (F) from control plants was extracted. The integrity of the RNA from shoots (B) and roots (C) of the 4 plants under the 4 water treatments, control (Ct), moderate water deficit (MWD), severe water deficit (SWD) and recovery (Rec), are shown.

Supplementary Table 1: *A. thaliana* Argonaute, Dicer-like and genes responsive to water deficit mRNAs and protein sequences used for identification of the genes in *M. truncatula*. Are indicated in the table the gene abbreviation names, The Arabidopsis Information Resource (TAIR) accession number of the genes are shown and their mRNA and protein accession numbers in NCBI.

Gene name	Gene Abbreviation	TAIR	mRNA accession number	Protein accession number
Dicer-like				
Dicer-like 1	DCL1	AT1G01040	NM_099986.3	NP_171612.1
Dicer-like 2	DCL2	AT3G03300	NM_111200.5	NP_566199.4
Dicer-like 3	DCL3	AT3G43920	NM_114260.1	NP_189978.1
Dicer-like 4	DCL4	AT5G20320	NM_122039.4	NP_197532.3
Argonaute				
Argonaute 1	AGO1	AT1G48410	NM_179453.2	NP_849784.1
Argonaute 2	AGO2	AT1G31280	NM_102866.2	NP_174413.2
Argonaute 3	AGO3	AT1G31290	NM_102867.1	NP_174414.1
Argonaute 4	AGO4	AT2G27040	NM_128262.3	NP_565633.1
Argonaute 5	AGO5	AT2G27880	NM_179779.2	NP_850110.1
Argonaute 6	AGO6	AT2G32940	NM_128854.3	NP_180853.2
Argonaute 7	AGO7	AT1G69440	NM_105611.3	NP_177103.1
Argonaute 8	AGO8	AT5G21030	NM_122111.2	NP_197602.2
Argonaute 9	AGO9	AT5G21150	NM_122122.2	NP_197613.2
Argonaute 10	AGO10	AT5G43810	NM_123748.2	NP_199194.1
Genes responsive to water stress				
Early Responsive to Dehydration 1	ERD1	AT5G51070	NM_124486.2	NP_568750.1
Plasma Membrane Intrinsic Protein 1;5	PIP1;5	AT4G23400	NM_118469.3	NP_194071.1

Supplementary Table 2: Primers used for gene quantification by quantitative Real Time PCR in *M. truncatula*. Indication of the amplification product size (Amplicon size) and the PCR efficiency used for each pair of primers obtained from real-time PCR Miner software (version 2.2). The efficiency for each gene was calculated doing the arithmetic mean of all efficiencies given by PCR Miner.

Gene	Side	Primer sequence	Amplicon size	Efficiency
Genes used to do the screen for selection of the qRT-PCR reference(s) gene(s)				
Actin 11	Forward	TTGGCATCACACCTTCTAC	158	0.81825
	Reverse	CCTGAATAGCAACATACATAGC		
Actin 7	Forward	GACAATGGAAGTGAATGG	196	0.8726
	Reverse	CAATACCGTGCTCAATGG		
Aprt	Forward	ATGTGCTGCGATTAAACTAC	139	0.9005
	Reverse	GGCTCCTTCTCCTTCAAC		
Elf-1a	Forward	ATTACCATTGATATTGCTTTGTG	186	0.87855
	Reverse	CTGTCCATCCTTAGAGATACC		
L2	Forward	GCTTACCACAAGTTCAGAG	164	0.90345
	Reverse	GCAATGAGACCAACCTTC		
HDT3	Forward	GCCAAGCACAGTGCCAAG	123	0.97335
	Reverse	CCATCAGCCAGTTTTCTAAATC		
Genes responsive to water stress				
ERD1	Forward	TTGGATTGATGAGCCGAGTG	112	0.9582
	Reverse	TGCAGCCTTTATGGCATCTT		
PIP1;5	Forward	ACTGGAAGTGGCATTAAACCC	83	0.97465
	Reverse	ATCCAGTGGTCATTCCAAGC		
Dehydrin	Forward	ACGGTGCTGTTCTATAAGC	115	0.87435
	Reverse	TCAACTGCGGTGGTTGTATC		
MI-1-P	Forward	AATTTAAGGAAGCAAACAAAG	98	0.81845
	Reverse	CATGGTGTGATTGAGTCC		
Dicer-like genes				
DCL1	Forward	CCATAAAGTGCAGAGAATGTAGG	160	0.90525
	Reverse	AAGACAAGAGACGCCTTCG		
DCL2	Forward	AGACTCGGTGCTGGACTATC	170	0.91165
	Reverse	GCTCTTGAGAGGCATGAAGG		
DCL3	Forward	ATATGCGGTGAGCTGTGCATC	116	0.9348
	Reverse	GATCTGTTCCAGCTTGTGT		
Argonaute genes				
AGO1	Forward	GTGGAAGAGGTGGATATAGC	134	0.8695
	Reverse	GTAGGGAGCAGAAGATACAG		
AGO12b	Forward	GGTTACCAACCTCCGGTTAC	104	0.9793
	Reverse	TTCCCACCTCTATCCATGCT		
Ago-07	Forward	CTGCCTTCTTCTCAATTACC	137	0.95545
	Reverse	GCCAGTTCATGCTACCAAC		
Ago-11	Forward	GCCTTGAGAGGTCAGGAATC	175	
	Reverse	GGAAGCCACGACAACATTG		
AGO4b	Forward	AAGCGACCTACATTTGTCCC	89	0.9733
	Reverse	GAGGATCTCTGGAGTGTTGGA		
AGO4c	Forward	ACAAGCAGACCAACACATTATC	176	0.89955
	Reverse	ATGAACTGTCCAACCTGAGATG		



Supplementary Figure 2: Staining of acrylamide the gels used to make the northern blot membranes. Gel stained with Sybr Safe (4ul to 50ml 1X MOPS) during 10min before and during 20min after the transferring of the RNA to a Hybond-NX membrane.(A) Gel with RNA from different tissues of *M. truncatula*: shoots (S) and roots (R), seedlings (Sd), seed pods (Sp) and flowers (F). (B) The RNA from the plants under the 4 water treatments, control (Ct), moderate water deficit (MWD), severe water deficit (SWD) and recovery (Rec).

13.2. Part 1 - Supplementary Tables and Figures

Supplementary Table 3: Functions of the genes used to determine the best reference gene(s) for the water deficit treatments imposed to *M. truncatula*.

Gene Name	Abreviation	Gene Function	Reference
60S Ribosomal Protein L2	L2 or RPL2	The ribosomal protein L2 is an essential component of the ribosomal large subunit (60S) and have a function in the subunit association.	Marty and Meyer, 1992
Adenine Phosphoribosyl Transferase 1	APRT	Adenine phosphoribosyl transferase (APRT) is the key enzyme that converts adenine and 5-phosphoribosyl-1-pyrophosphate to adenosine monophosphate.	Moffatt et al. 1991
Actin 11	ACT11	ACT11 increase its expression in the ovule, embryo, and endosperm, indicating that it plays a role in the reproductive life cycle of Arabidopsis.	Huang et al, 1997
Actin 7	ACT7	Arabidopsis ACT7 genes involved in the regulation of hormone-induced (auxin and cytokinin or auxin alone) plant cell proliferation and callus formation.	Kandasamy et al, 2001
Elongation Factor 1-Alpha	ELF-1 α	ELF-1a is a subunit of the protein synthesis factor eFEL and binds the aminoacyl-tRNAs to the acceptor site of ribosomes during peptide chain elongation.	Sun et al, 1997
Histone Deacetylase 3	HDA3	The histone deacetylases regulates the status of histone acetylation in a cell and HDA3 are required to establish leaf polarity.	Ueno et al, 2007

Supplementary Table 4: Some characteristics of the genes used in the screen of the best reference gene(s) in the water deficit treatments imposed to *M. truncatula*. ORF, Open Reading Frame; MTGI, *Medicago truncatula* Gene Index (MTGI) 9.0 reference; Mt2.0 genecall, *Medicago truncatula* Genome Sequencing Project release version 2.0 (Mt2.0)accession number; Crom., chromosome localization.

Gene	Mt2.0 genecall	ORF size (bp)	MTGI	Chrom.	Genomic region	Blastn and Blastp
L2	AC135311_31	780	TC133735	7	16423357-16421921	L2 (At2g18020)
APRT			TC114600			APRT (AT1G27450)
ACT11	AC184161_16	1134	TC119043	7	3709067-3711803	ACT11 (AT3G12110)
ACT7	AC137836_27	1149	TC124607	3	34688024-34686876	ACT7 (AT5G09810)
EF-1α			TC114384			EF-1a (AT1G07940)
HDA3			TC130795			HDA3 (AT3G44750)

Supplementary Table 5: Some characteristics of ERD1, PIP1;5, Dehydrin and MI-1-PS *M. truncatula* genes. ORF, Open Reading Frame; MTGI, *Medicago truncatula* Gene Index (MTGI) 9.0 reference; Mt2.0 genecall, *Medicago truncatula* Genome Sequencing Project release version 2.0 (Mt2.0 genecall) accession number. databases, Mt2.0; Chrom., chromosome localization.

Gene name	Abbreviation	Mt2.0 genecall	ORF size (bp)	MTGI	Chrom.	Genomic region	Blastn and Blastp
Early Responsive to Dehydration 1	ERD1	CT010504_34	2469	TC123036	3	27304045-27297848	ERD1 (AT5G51070)
Plasma Membrane Intrinsic Protein 1;5	PIP1;5	CT573507_13	864	TC117941	5	2379410-2377969	PIP1;5 (AT4G23400)
Dehydrin		AC141922_11		TC137420	4		Putative dehydrin (At1g76180)
Myo-inositol-1-Phosphate Synthase	MI-1-PS	AC125389_5	1553	TC115706	3	25226161-25229597	MI-1-PS (AT2G22240)

13.3. Part 2 - Supplementary Tables and Figures

Supplementary Table 6: Targets of miR162, miR168a, b and c.

miRNAs	TC or Genecall of the miRNA target	Gene	Gene Function	Cleavage site of the target validated
miR162	AC150443_32.2	Dicer Like 1	miRNA Biogenesis	Jagadeeswaran et al, 2009
miR168a	TC126820	Argonaute 1	miRNA Biogenesis	Jagadeeswaran et al, 2009
miR168b	TC126820	Argonaute 1	miRNA Biogenesis	Jagadeeswaran et al, 2009
miR168c	TC126820	Argonaute 1	miRNA Biogenesis	Jagadeeswaran et al, 2009
	TC140219	Unknown	Unknown	Not validated



Supplementary Figure 3: Cleavage site of the miR168a/b and miR168c in Argonaute 1 (AGO1) mRNA.

The arrow is pointing to the site of cleavage in the mRNA sequence between the nucleotide 10 and 11 (from 5' to 3') of miR168.

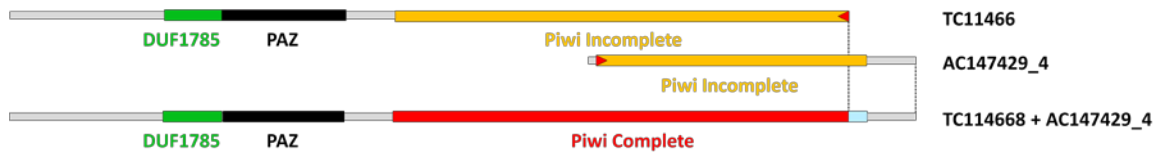
13.4. Part 3 - Supplementary Tables and Figures

Supplementary Table 7: Protein domains present in the translated TC114668 sequence, in protein AC147429_4 sequence and in the sequence originated from the joining of both sequences (TC114668 + AC147429_4). The protein domains were obtained using the Conserved Domains Database (CDD) database of NCBI.

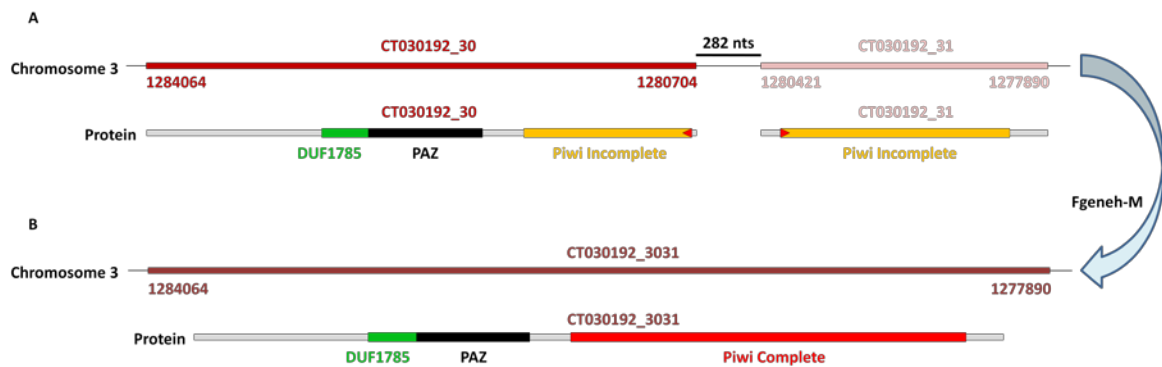
Gene Name	Mt2.0 Genecall	MTGI 9.0	Protein Domains (CDD)		
			DUF1785	PAZ	PIWI
MtAGO4a	AC147429_4	AC147429_4	142-194	194-306	351-763
MtAGO4a		TC114668			8-253
MtAGO4a	TC114668 + AC147429_4		142-194	194-306	351-779

Supplementary Table 8: Protein domains from CT030192_30, CT030192_31 and the sequence originated from the joining of both protein sequences (CT030192_3031). The protein domains were obtained using the Conserved Domains Database (CDD) database of NCBI.

Gene Name	Mt2.0 Genecall	Protein Domains (CDD)		
		DUF1785	PAZ	PIWI
MtAGO11	CT030192_30	192-244	244-368	413-596
MtAGO11	CT030192_31			22-273
MtAGO11	CT030192_3031	192-244	244-368	413-846



Supplementary Figure 4: Scheme demonstrating the TC114668 and AC147429_4 protein common regions, and the fusion done with both sequences. TC114668 and AC147429_4 give two incomplete proteins that complement each other, that after the junction it originates one complete protein. The protein domains were obtained using the Conserved Domains Database (CDD) database of NCBI.



The images aren't at scale

Supplementary Figure 5: Scheme showing the genomic region of CT030192_30, CT030192_31 in chromosome 3 and the distance between both sequences. CT030192_30 and CT030192_31 give two incomplete proteins. But after use of Fgenesh-M software in the genomic region belonging to both genes, was obtained just one coding sequence originating one complete protein. The protein domains were obtained using the Conserved Domains Database (CDD) database of NCBI.

Supplementary Figure 6: Nucleotide coding sequence of AGO4a resulted from the junction of Black TC114668 (black) and AC147929_4 (blue).

```

CCCTGTGGAAGGGAAGGGTGTGGGAGGAAGATTATGGACAGGGTGCAGGAGACATACGCCCTCGACTTGAATGGCAAGGAATTTGCATATGATGGGGAGAAAAGCT
TGTTTACTATTGGCTCTCTCCACAAAACAAGCTCGAGTTTGAATTTGTTCTAGAGGATGTTGTTTCCAGCAGAAAATATGGAATCGTAGCTCGTAGCTAATGGG
GACAATGAAGCTGACAAAAGAGGGTGCAGGCCATACAATGCAAAGACATTTAAAGTTGAGATTAGCTTTGCTACAAAAGTTCCTATGTATGCCATAGCCAATGC
GTTACCGGGCAGGAAACTGAGAAATTTCCAAGAAGCAGTCAGAGTTCTCGATATCATATTTAGGCCAACATGCTGCTAAGCAGGGCTGCCTACTTGTGGCCAAATCTT
TCTTCCACAATGATCCCAAGAATTTTGTCTGATGTGGGGGGTGGTGTCTTAGGCTGTCGAGGATTCATTCAAAGTTTATAGAGCTACACAGAGCGGGCTCTCTTAAC
ATAGATGTCCTCAACTACCATGATAATTCACCTGGGCCTGTGGTGGATTCTTAATTTCCAATCAAATGTTAGAGATCCTTTTCAGATTGACTGGGGAAAAGGCCAA
GAGGACCCATAAAAATCTGAGGGTGAAGACTCATCCATCCAATCAAGAGTGGAAAATATGCGGACTCAGCGAAGTCCCATGCAAAGAGCTTACTTTTACTTTGAAGA
AAAGGGATGGTGTACTGATGAGATGACTGTTCTGGATTTTACTAATGTCCGGAAGATAGATCTGCGCTACTCTGCTGATCTCCATGATTAATGTTGGC
AGGCCAAAACGGCCAACATATTTCCCATTTGAGCTGTGTGAATGGTATCACTCCAACGATATACAAAAGCTTTCCACACTTCAAAGGGCTTCACCTGGTGGAGAA
GTCCAGACAGAAGCCACAAGAGAGAATGAGAATTTTGTCCGATGCCTTAAAATAGCAACTACGGGGCTGAACCTCTGCTTCAAAGTTGTGGAATTTCTATAAGCA
CAGGCTTCACTCAAGTGGAGGACGTGTTTTCCTGCTCCAAAGTTGAAGTTCGGCAATGGTGAAGATTTCACTCCGAGAAAATGGAAGATGGAATTTCAACAACAAG
AAATTTGTGCAGCCAAACAAGAAGATAGAAAATGGGCTGTGGCAAACCTTTCTGCACGATGTGATGTACGAGGACTTGTGAGGGATATAATTAGAATTTGAAAATAT
GAAAGGAATATGATGACCAACCATTCGATGTTTGAAGAAAATCCTCAGTTTAGACGTGCTCCACCAATGGTGAAGGATAGAGAAAATGTTGAGGACATTCAGT
CCAAACTCCAGGGGCTCCTCAATTCCTTCTTTGTTTGTCTGATCGGAAAACCTGTGATATTTATGGCCATGGAAAAGGAAGAAATCTTGCTGATTTTGAATC
GTTAATCAGTGCATGTCTCTTTGAGACTCAATGATAATTTATGGGTAATATATGTTGAAGATAAATGCCAAGCTTGTGGTGAAGTCTTTGTTGGTGTGTTGA
AAGTCTCCATCTCTTCTATTGTTTCCAAAGCACCTACCTCATTCTAGGAATGGATGTGTCCATGGCTCACCGGGCAGACTGATAITTCCTCAATTTGCCGCGG
TTGTGAGCTCTAGACAGTGGCCTCTGATTTCAAATATAGGGCATGTGTTCCGACACAGTCTGCAAAAAGTTGAAATGATAGATAAATTTGTTCAAGAAAGTATCAGAT
ACCGAGGATGAAGGCATCATGAGGGAACCTTGTCTCGATTTCTATACTAGTTCTAAGAATAGAAAACAGATAACATAATTTATTTCAAGGATGGTGTAGTGAGTC
ACAATTCATCAAGTTTGAATATTGAACCTCGACCAAGATCAITGAGGCATGCAAGTTCTTGTGATAAATTTGACTCCAAAATTTGTTGTAATTTGTTGCTCAAAAAG
ACCACCACACAAGATTTCCAGCCCAATTTCTCCTGACAACGTCACCACTGGAAGTATTTATGCAATAAAAATTTGTCATCAAAGAACTATGATTTCTACCTCTGT
GCTCATGCTGGATGATAGGCCACTAGGCTACCCACTACCCTACCTGCTGCTGATGATGAGATTGGCTTCTCCCTGATGAACCTCAGGAGCTTGGCTATTCCCTGCT
ATATGTTTACCAGAGGAGCAGCACTGCTATTTTCAGTTGTTGCACTTATATGCTATGACACTTGGCCGCTACTCAGTTGGGGAGTTCAIGAAAATTCGAGGATAAAT
CTGAGACATCTCAAGTCATGGTGGGCTGAGTGTGTCAGGAGCTGTACCTGTCCCTCAGCTGCCAAAAGTTCAGGACAACCGTGTGCAACTCCATGTTCTTTGT
TTGA

```

Supplementary Figure 7: Nucleotide coding sequence of AGO11 resulted from the use of Fgeneh-M software in the genomic region belonging to both genes, obtaining a coding region constituted for the junction of CT030192_30 (black) and CT030192_31 (blue).

ATGGAATACTCATTGAACAAATGAGGAAATGAGAAAGACATTGAATTCAAACCATGTGCCCATGGCCAGGAGAGGTCTTGGATCCAAGGGAGCCAAAGATACAACCTT
 AGCGAACCATTTTCAGATTGGCCTGAGCAAGAATGACGGCTACTTCTATCACTACAATGTTGCTTTATGTTATCAAGATGGACATGCGGTTGAAAGTAAAGGAGTTG
 GGAGAAAAGTCAATGACAAGCTTTGTAACAATATGATGTATTAAGAAACAAGAAATTTGCATATGATGGCGAGAAAAGTTATCTACTCTTCGTTCTACATCAT
 AAAAAAGCAGGAATTCATTGTTGTTAGAGGAAGTTTCATCAACGAGAGTTGGAAGCAACCCCTTCTGAAGCAACCAAGAGAATGAAGCATCAATCTCGGTCCAAAAC
 ATTTAAGGTTGAGATTAGTCATGTATCTAAAAATTCGGTTACAAGAAATACAGATGCCITGAGAGGTCAGGAATCTGAACACTATCAGGAAGCATTAAATTTCCCTG
 ACACAATACTGAGGCAGAAATGACGCAAAAACAGGGATGCTCCGATATTCATAAATCGTACTTTACAGATAATCAAAAAGAAATTTACAACCTGGAAAGTGGTATTCAA
 TGTTCCTCGTGGCTTCCATTCCGAGCTTCAGAGTACACAAAAGAGGTTTATCTCAATGTTGATGTGCAACTACATTAATGTAAGCCTGGTCTGTAGTGGACTT
 CCTTCTCCAGAATCAGAATGTTGAGAAACCAAACTTGATTGATTGGACCAGGTTATATTGCTCCTCCATCTCGAAGTAGAGGCAAAAAGAAATGTTGAAAAATCTTA
 GAATCAAGGCTAAACAATACACAACGCAAAATCACTGGACTAAGTGAAGTCCCTGCATGACACAAAATTTTTTTTAAACATGGAAATGATGCTAATGGTGAAGTG
 CAGTCAAGTGAATAACAATTTATGATACTTCAAACGCCACAAAAGATTGAACCTTGTATTCTGTTGATATGCCATGCATCAATGTTGAAAACCAAGAGGCC
 AATTTATTACCCTATGGAGTTATGCACATTAGTCTCATTGACGCGATACACTAAGCCGTTAGCGCATAAAACAAGGGCTCAACTGATATGGAAATCAAGGACTAGTC
 CTCGTGAAGGAAAGAAAGCTTTGCAATATTTCTTAAAGAAACACGAGATATGGCGATGAGCCATGCTTCGCTCTTTAGGGATTACTATCGAACCTTCCCTTACACAA
 GTTGACCGTCTGCTTTCGCAACTCCGACACTTATTGTTGGAAGGGGACAGAATTTTGTCCCGAAATGGGAGCTGGAATTTCAATGATAAGAAAATTAATTTGAACC
 TGTGAAGATTAAGCGTTGGCAATTTGAAACTTTTCGTCACAGTGTGATACGAAACATTTGTGCAGTATGATTAATAAATGTTCTGAAATGAAAGGATGCTCATAG
 ATCCTCCGTTCCGACATATTGAAAGGATATTGCGCATAGGAATGAGTGCCTTTCGCGAGGTTGCAAGAATGATGAAATGTAAGGCAAGCTACCTGGAGCC
 CCTACCGATCCTCTAGCTCAGTCTTTTGTGTATTCTCCGTTAGCAGGAATTTAACAATTTATGGTCTTGGAAAAGAAAGATGCTTGTGATGAAGGATTC
 AACACAATGCATTGCTCCACAAAATTAATGATCATTACATTATTAATGTTCTATTGAAAATTAATGCAAAAGCTGGTGGGATGAATTCCTTTTGTAACTGAAT
 TTAAGCATTTCGATACCCTGTTTCAAAAATACCTACATTTGGTTATTTGGAATGGATGTTTCACATGGATCTCAAGGTCATCAGAAGCACTATCCATTGCTCGCGTT
 CTGAGCTCAAGATGTTGGCTCAAACTCTCGTTACAAGCTGTAGTCCGCACTCAGTCCGTCAAAAGTAGAGATCGTTCACTCTCTAATTCAGGCTGTTTCTGATAC
 TAAGGATGACGGTATCATCAGTGTGCTTAAAGACTTCCAACAACCTCAGGAGTAAAGCCCTCAACAAATTAATAATTTTCAGGAGCGGCGTGAAGCAATCACAGT
 TCAACCAAGTCCCTCAACATTGAAATGAAATCATAAAGGCATGTAAGTGTATGATGAAAGTGGTGTCCAAAGTTTACTTTGATTGTTGCTCAGAAGAATCAT
 CACACTAGTCTTCAAGGCAATTCCTCAAGAAAATTTTACCTGGAACCTGTTATTTGATAACACTATATGTCATCCCAAAGCAATGACTTCTACATGTGTGC
 ACATGCTGGGAGGATGGCACAAGTGTCTACTCATTACCATGTTTATATGATGAAATGGGTTCTCAGCAGATAAATTTACAAGAAATTTGTGCATCTCTATGTT
 ATGTGACCCAGAGAAGCAAAATGCCATATCAATAGTTGCACCAATATATATGCTGACCTAGCTGCTGCTCAGATAGCACAATTTATAAAAATATGATGAGTCAGAG
 AATCTTTCCAGCCACAATGAAATTCATATCTCAAAATACCTACTGAGTTACCACGCTTCTCATGAACGAGTTGCAGATTCATGTTCTCTGTTAA

Supplementary Table 9: Protein domains and their localization on the Dicer-like protein sequences of *M. truncatula* using the Conserved Domains Database (CDD) database of NCBI.

Gene Name	Protein Domains (CDD)							
	DExD	Helicase-c	Duf283	PAZ	RNasellI-a	RNasellI-b	dsRB-a	dsRB-b
MtDCL1	234-386	624-740	816-905	1149-1354	1399-1582	1617-1772	1776-1836	1865-1933
MtDCL2	33-187	383-483	555-628	806-917	964-1115	1147-1301	1305-1370	
MtDCL3	67-214	398-521	585-668	917-1053	1096-1260	1306-1456	1462-1528	1609-1692

Supplementary Table 10: Protein domains and their localization on the Argonaute protein sequences of *M. truncatula* using the Conserved Domains Database (CDD) database of NCBI.

Gene Name	Protein Domains (CDD)		
	DUF1785	PAZ	PIWI
MtAGO1	324-376	376-490	535-986
MtAGO12a	178-230	230-344	389-839
MtAGO12b	204-241	241-255	400-490
MtAGO2	17-69	71-185	233-669
MtAGO7	316-368	379-484	533-976
MtAGO4a	142-194	194-306	351-779
MtAGO4b	222-274	274-390	435-867
MtAGO4c	188-240	240-357	401-831
MtAGO11	192-244	244-368	413-846

Supplementary Figure 8: Amino acid alignment of the PIWI domains of *M. truncatula* (Mt) and *A. thaliana* (At) Argonaute proteins. The protein sequences were aligned using CLUSTAW. The amino acids residues corresponding to the conserved aspartate, aspartate and histidine (DDH) triad residues are marked in red, while the *A. thaliana* Argonaute 1 histidine in the position 800 (H800) is in blue. Amino acid positions corresponding to the beginning and end of the PIWI Domains in each protein are mentioned.

AtAG04	626	KI NAKLGGNLSM ^{SVERTPAFT} - - VI SKVPTI I LGMDVSHGSPGQSD- VPSI AAVVSSREWPLI SKYRASVRTQPSKAEMI ESLVKKNG- TE - - - DDGI I KELLVDFYT
MtAG04a	525	KI NAKLGGNLSLLGVESSPSLP - - I VSKAPTLL LGMDVSHGSPGQTD- I PSI AAVVSSRQWPLI SKYRACVRTQSAKVEMI DNLFKKVSDE - - - DEGI MRELLLDFT
MtAG04b	572	KI NAKLGGNLSLLGVEHSPSP - - I VSKAPTLL LGMDVSHGSPGQTE- I PSI AAVVSSRQWPLI SKYRACVRTQSAKVEMI DNLFKKVSDE - - - DEGI I RELLDLDFYN
MtAG04c	608	KI NAKLGGNLSWLVGHSRSI P - - I VSKVPTLI LGMDVSHGSPGQPD- I PSI AAVVSSRQWPLI SKYRACVRTQSAKVEMI DNLFKKVSDE - - - DEGI I RELLDLDFH
AtAG08	570	KI NAKLGGNLSVLDMEISGTM - - LVMRPTI I I GMDVSHGSPGQSDHI PSI AAVVSSREWPLI SKYRACVRTQSAKVEMI DSLFKKVSDE - - - DQGI MRELLLDFFH
AtAG09	598	KI NAKLGGNLSLLAMERSPAMP - - KVTQVPTI I VGMVSHGSPGQSD- I PSI AAVVSSRQWPLI SKYKACVRTQSAKMEMI DNLFKKVSDE - - - E - GMFRELLEDFY
MtAG011	591	KI NAKLGGNLSLLTEFKHSI P - - LFPSK PTLVI GMDVSHGSPGQSD- ALSI AAVVSSRQWPLI SRYKAVVRTQSAKVEI VQSLFKKVSDE - - - DDGI I SELLDKDFQT
AtAG06	589	KI NSKLGGI NSLLGI EYSYNI P - - LI NKI PTLI LGMDVSHGSPGQSD- VPSIAAVVSSRQWPLI SRYRAAVRTQSPRELEI DSLFKKVSDE - - - GDNGI MRELLDFEYR
AtAG05	687	KI NVKTTGGNRTVLDVDAI SCRIP - - LI TDRPTI I MGADVTHPPGQEDS- SPSI AAVVASMDWPEI NKYRGLVSAQAAREE I QDLKYLQDPQRLVHSGLI REHFI AFRR
MtAG012a	569	KI NVKMGGRNTVLDVDAI RRPI - - LVSVDPTI I FGADVSHPESEDV- CPSI AAVVASQDWPEVTKYAGLVCAQAPREI I KDLFKKVSDE - - - DDGI I SELLDKDFQT
MtAG012b	575	KI NVKMGGRNTVLDVDAI SCRIP - - LVSVDPTI I FGADVTHPEGEES- SPSI AAVVASQDWPEVTKYAGLVCAQAPREI I QDLKYLQDPVVRGTVSGGMI RDLII SFRK
AtAG010	728	KI NVKMGGRNTVLDVDAI RRPI - - LVSVDPTI I FGADVTHHPGEGDS- SPSI AAVVASQDWPEI TKYAGLVCAQAPREI I QDLFKKVSDE - - - DDGI I SELLDKDFQT
MtAG01	716	KI NVKMGGRNTVLDVDAI RRPI - - LVSVDPTI I FGADVTHHPGEGDS- SPSI AAVVASQDWPEI TKYAGLVCAQAPREI I QDLFKKVSDE - - - DDGI I SELLDKDFQT
AtAG07	713	KI NAKI GGSMTALYNSI PSHI PRLLRPDEPVI FGADVTHHPGEGDCC- SPSIAAVVGSINWPEANRYVSRMSQTTRQEI I QDL - - - LMVKELLDFFYK
MtAG07	698	KI NAKVGGCTVALYNSLPRLFNIDEPVFMFGADVTHHPGLDSS- SPSIAAVVGSINWPEANRYVSRMSQTTRQEI I ADLG - - - AMWGELLEDYFK
MtAG02	420	KI NAKI GGSNVELI N - RLPHFE - - - DESHVMFI GADVNHGSRDITN- SPSI VAVVATTNWPANRYAARVCAQAEICTEKI LNFG - - - EICLDLVRHYEK
AtAG02	823	KMNAKVGGSNVELMD - TFSFFK - - - KEDEVMI GADVNHAAARDKM- SPSI VAVVGTLNWPANRYAARVIAQPRKEEIQCFG - - - DACLELVKAHYK
AtAG03	715	KI NAKVGGCTVALYNSI PSHI PRLLRPDEPVI FGADVTHHPGEGDCC- SPSIAAVVGSINWPEANRYAARVCAQAEICTEKI LNFG - - - EICLDLVRHYEK
AtAG04		SSNKRKPEHI I FRDGVSEQFNQVLI ELDQI I EACKLLDANWPKFLLVAQKNHHTKFFQPTSP - - - - - ENVPPTGI I DNKI CHPKNDFYLCAHAGMI G
MtAG04a		SSNKRKPDNI I FRDGVSEQFNQVLI ELDQI I EACKFLDENWTPKFVVI VAQKNHHTKFFQPNSP - - - - - DNVPPTGI I DNKI CHPRNDFYLCAHAGMI G
MtAG04b		SSNKRKPDNI I FRDGVSEQFNQVLI ELDQI I EACKFLDENWTPKFVVI VAQKNHHTKFFQPNSP - - - - - DNVPPTGI I DNKI CHPRNDFYLCAHAGMI G
MtAG04c		SSNKRKPDNI I FRDGVSEQFNQVLI ELDQI I EACKFLDENWTPKFVVI VAQKNHHTKFFQPNSP - - - - - DNVPPTGI I DNKI CHPRNDFYLCAHAGMI G
AtAG08		SS - GKPNHI I FRDGVSEQFNQVLI ELDQMI - - - - - I NHHTKFFQTESP - - - - - NNVLPGTI I DSNI CHQHNDFYLCAHAGMI G
AtAG09		SSNKRKPEHI I FRDGVSEQFNQVLI ELDQMI - - - - - I NHHTKFFQTESP - - - - - NNVLPGTI I DSNI CHQHNDFYLCAHAGMI G
MtAG011		TS - GVRPQI I FRDGVSEQFNQVLI ELNEI I KACQCYDESWCPKFTLI VAQKNHHTKFFQPNSP - - - - - ENVPPTGI I DNKI CHPKNDFYLCAHAGMI G
AtAG06		TSRARKPKQI I FRDGVSEQFNQVLI ELDQI I KAYQRLGESDVPKFTVI VAQKNHHTKFFQPNSP - - - - - ENVPPTGI I DNKI CHPRNDFYLCAHAGMI G
AtAG05		AT - GQIPQRI I FYRDGVSEGFQVLLHEMTAI RKACNSLQENYVPRVTFVI VQKRHHTLFPQHQNRDMTD - - - - - KSGNI QPPTVVDTKI CHPNEFDYLCASHAGI QG
MtAG012a		AT - GKKPCRI I FYRDGVSEGFQVLLHEMTAI RKACNSLQENYVPRVTFVI VQKRHHTLFPQHQNRDMTD - - - - - KSGNI QPPTVVDTKI CHPNEFDYLCASHAGI QG
MtAG012b		AT - GKKPCRI I FYRDGVSEGFQVLLHEMTAI RKACNSLQENYVPRVTFVI VQKRHHTLFPQHQNRDMTD - - - - - KSGNI QPPTVVDTKI CHPNEFDYLCASHAGI QG
AtAG010		AT - GQKPLRI I FYRDGVSEGFQVLLHEMTAI RKACASLEPNYQPPVTFVI VQKRHHTLFPQHQNRDMTD - - - - - RSGNI LPPTVVDTKI CHPTEFDYLCASHAGI QG
AtAG01		ST - GHKPLRI I FYRDGVSEGFQVLLHEMTAI RKACASLEAGYQPPVTFVIVQKRHHTLFPQHQNRDMTD - - - - - RSGNI LPPTVVDTKI CHPTEFDYLCASHAGI QG
MtAG01		AT - GQKPCRI I FYRDGVSEGFQVLLHEMTAI RKACASLEPNYQPPVTFVIVQKRHHTLFPQHQNRDMTD - - - - - RSGNI LPPTVVDTKI CHPTEFDYLCASHAGI QG
AtAG07		AV - KKLPNRI I FRDGVSEGFQVLLHEMTAI RKACASLEPNYQPPVTFVIVQKRHHTLFPQHQNRDMTD - - - - - HENI PPTVVDTKI THPKFDFYLCASHAGI QG
MtAG07		EV - EKLPNRI I FRDGVSEGFQVLLHEMTAI RKACASLEPNYQPPVTFVIVQKRHHTLFPQHQNRDMTD - - - - - HENI PPTVVDTKI THPKFDFYLCASHAGI QG
MtAG02		LN - KVRPQI VI FRDGVSEGFQVLLHEMTAI RKACASLEPNYQPPVTFVIVQKRHHTLFPQHQNRDMTD - - - - - GNVFPTVVDTKI THPKFDFYLCASHAGI QG
AtAG02		AT - GKRPNKI VI FRDGVSDAQFDMVNLVLDVCLTFE - - - - - KNGYNPKI TVI VAQKRHHTLFPQHQNRDMTD - - - - - GNVFPTVVDTKI THPKFDFYLCASHAGI QG
AtAG03		AP - EKRPNI VI FRDGVSDAQFDMVNLVLDVCLTFE - - - - - KNGYNPKI TVI VAQKRHHTLFPQHQNRDMTD - - - - - GNVFPTVVDTKI THPKFDFYLCASHAGI QG
AtAG04		TTRPPTHYHVLDEI GFSADELQELVHSLSYVQRSTTAI SVVAPI CYAHLAAQA - - - - - 879
MtAG04a		TSRPTHYHVLDEI GFSPDELQELGHSLSYVQRSTTAI SVVAPI CYAHLAATQ - - - - - 779
MtAG04b		TSRPTHYHVLDEI GFSPDDLQELVHSLSYVQRSTTAI SVVAPI CYAHLAASQVGFQM 867
MtAG04c		TSRPTHYHVLDEI GFSPDDLQELVHSLSYVQRSTTAI SVVAPI CYAHLAASQVGFQM 831
AtAG08		TTRPPTHYHVLDEI GFSDTLQELVHSLSYVQRSTTAI SVVAPI CYAHLAAQA - - - - - 805
AtAG09		TTRPPTHYHVLDEI GFATDDLQELVHSLSYVQRSTTAI SVVAPVCYAHLLAAQA - - - - - 851
MtAG011		TSRPTHYHVLDEI GFSADELQELVHSLSYVQRSTTAI SVVAPI CYAHLAAQA - - - - - 846
AtAG06		TSRPTHYHVLDEI GFSPDDLQELVHSLSYVQRSTTAI SVVAPVCYAHLLAAQA - - - - - 850
AtAG05		TSRPAHYHVLDEI GFTADLQELVHSLSYVQRSTTAI SVVAPVCYAHLLAAQA - - - - - 957
MtAG012a		TSKPAHYHVI WDNKFSADEI QSLTNNLCYTYARCTRSVLPVPPYAHLLAAYRARFYM 839
MtAG012b		TSKPAHYHVI WDNKFSADEI QSLTNNLCYTYARCTRSVLPVPPYAHLLAAYRARFYM 839
AtAG010		TSRPAHYHVLWDENNFTADGI QSLTNNLCYTYARCTRSVLPVPPYAHLLAAYRARFYM 945
AtAG01		TSRPAHYHVLWDENNFTADGLQSLTNNLCYTYARCTRSVLPVPPYAHLLAAYRARFYM 998
MtAG01		TSRPAHYHVLWDENNFSDGLQSLTNNLCYTYARCTRSVLPVPPYAHLLAAYRARFYM 986
AtAG07		TSRPTHYHVLWDENNFTSDELQELVHSLSYVQRSTTAI SVVAPVCYAHLLAAYRARLYI 949
MtAG07		TSRPTHYHVLWDENNFTSDELQELVHSLSYVQRSTTAI SVVAPVCYAHLLAAYRARLYI 976
MtAG02		TSKPTHYHVLWDEHRTSDELQELVHSLSYVQRSTTAI SVVAPVCYAHLLAAYRARLYI 669
AtAG02		TSKPTHYHVLWDELGFTSDVQKLI YDMCFTRCTKPVSLVPPVYAHLLAAYRARLYI 963
AtAG03		TSKPTHYHVLWDELGFTSDVQKLI YDMCFTRCTKPVSLVPPVYAHLLAAYRARLYI 1074

Acknowledgements

I start to show gratitude first to my two coordinators, Professor Manuel Pedro Salema Fevereiro (PhD) and the PhD Dulce Maria Metelo Fernandes dos Santos. On especial thank to Jorge Paiva that supported me with the qRT-PCR technique, to Inês Trindade that helped me with the northern blot technique and to achieve the status of “Dr. Capitão” in the lab and to Susana for her advises for my thesis and for call me “piqueno”. I want to show appreciation too for Mara, Nuno, Leonor, Silvana, Carlota, Susana and Rita that had the patience for my mistakes (not very severe, thank God!) and to explain me several thinks in this time that I spent in the lab. And finally show my gratitude to my mother’s and to brother’s support, to my brother Luciano and to Mariana just for all.

Thanks to all of you