



# Small RNA pathways and diversity in model legumes: lessons from genomics

Pilar Bustos-Sanmamed<sup>1</sup>, Jérémie Bazin<sup>1,2</sup>, Caroline Hartmann<sup>1,2</sup>, Martin Crespi<sup>1\*</sup> and Christine Lelandais-Brière<sup>1,2</sup>

<sup>1</sup> Centre National de la Recherche Scientifique, Institut des Sciences du Végétal, Gif-sur-Yvette Cedex, France

<sup>2</sup> Université Paris Diderot, U.F.R. Sciences du Vivant, Paris Cedex 13, France

## Edited by:

Georgina Hernandez, Universidad Nacional Autónoma de México, México

## Reviewed by:

Ramanjulu Sunkar, Oklahoma State University, USA  
Robert Stupar, University of Minnesota, USA

## \*Correspondence:

Martin Crespi, Centre National de la Recherche Scientifique, Institut des Sciences du Végétal, Bât 23, Avenue de la Terrasse, 91198 Gif sur Yvette, France  
e-mail: martin.crespi@isv.cnrs-gif.fr

Small non-coding RNAs (smRNA) participate in the regulation of development, cell differentiation, adaptation to environmental constraints and defense responses in plants. They negatively regulate gene expression by degrading specific mRNA targets, repressing their translation or modifying chromatin conformation through homologous interaction with target loci. MicroRNAs (miRNA) and short-interfering RNAs (siRNA) are generated from long double stranded RNA (dsRNA) that are cleaved into 20–24-nucleotide dsRNAs by RNase III proteins called DICERs (DCL). One strand of the duplex is then loaded onto effective complexes containing different ARGONAUTE (AGO) proteins. In this review, we explored smRNA diversity in model legumes and compiled available data from miRBase, the miRNA database, and from 22 reports of smRNA deep sequencing or miRNA identification genome-wide in three legumes: *Medicago truncatula*, soybean (*Glycine max*) and *Lotus japonicus*. In addition to conserved miRNAs present in other plant species, 229, 179, and 35 novel miRNA families were identified respectively in these 3 legumes, among which several seems legume-specific. New potential functions of several miRNAs in the legume-specific nodulation process are discussed. Furthermore, a new category of siRNA, the phased siRNAs, which seems to mainly regulate disease-resistance genes, was recently discovered in legumes. Despite that the genome sequence of model legumes are not yet fully completed, further analysis was performed by database mining of gene families and protein characteristics of DCLs and AGOs in these genomes. Although most components of the smRNA pathways are conserved, identifiable homologs of key smRNA players from non-legumes, like AGO10 or DCL4, could not yet be detected in *M. truncatula* available genomic and expressed sequence (EST) databases. In contrast to *Arabidopsis*, an important gene diversification was observed in the three legume models (for DCL2, AGO4, AGO2, and AGO10) or specifically in soybean for DCL1 and DCL4. Functional significance of these variant isoforms may reflect peculiarities of smRNA biogenesis and functions in legumes.

**Keywords:** small RNA, dicer, argonaute, model legumes

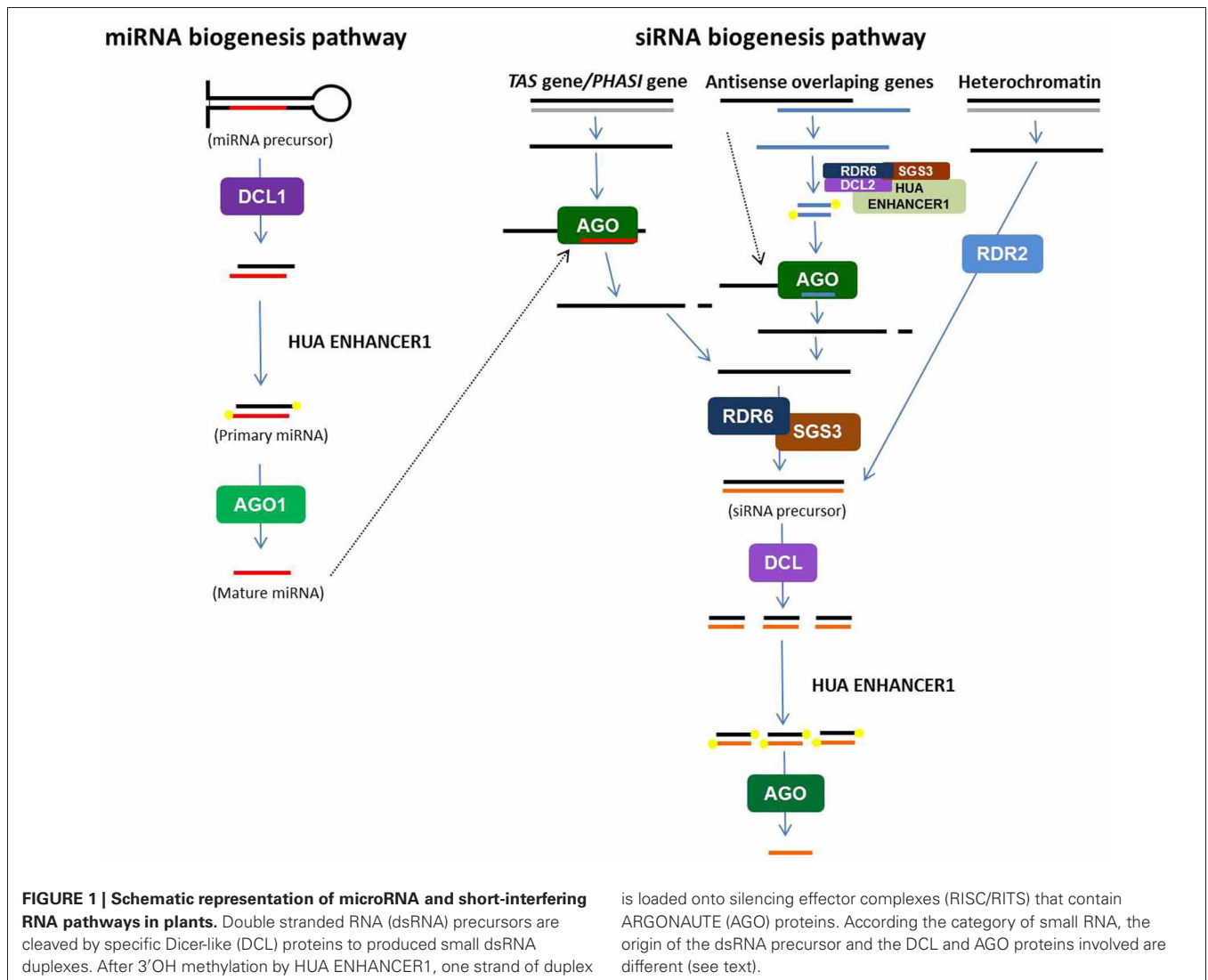
## INTRODUCTION

Small RNAs (smRNAs) are important riboregulators in bacteria, fungi, plants and animals, which negatively regulate the expression of specific target genes by base-pairing. In the last decade, smRNA functions have been largely described in plants, both in development and responses to biotic and abiotic interactions (for review, Khraiwesh et al., 2012). Plant smRNAs, 20–24 nucleotides (nt) in length, are classically divided into microRNAs (miRNAs) and short-interfering RNAs (siRNAs). Like miRNAs, the natural-antisense siRNAs (natsiRNA) and the trans-acting siRNAs (tasiRNAs) are two classes of siRNAs involved in post-transcriptional gene regulation by degrading and/or inhibiting translation of their mRNA targets. In contrast, the heterochromatin-associated siRNAs (hcsiRNA) are associated to chromatin modifications and transcriptional repression of their target DNA loci.

All plant smRNAs are generated from long double stranded RNA (dsRNA) precursors that are cleaved by RNase III proteins,

called DICER-like (DCLs). However, miRNAs and siRNAs derive from different types of precursors (Figure 1; Voinnet, 2009). The result of DCL action is the production of 20–24-base-pair (bp) RNA duplexes with 2 nt long 3' overhangs (Vazquez, 2006). After stabilization by 3'OH methylation (Bove et al., 2006), one strand of the smRNA duplex is loaded onto a silencing effector complex (called RISC/RITS), that contains one ARGONAUTE (AGO) protein, to mediate gene silencing by base pairing with their targets (Vazquez, 2006; Vaucheret, 2008; Mallory and Vaucheret, 2010). In *Arabidopsis thaliana*, 4 DCL and 10 AGO genes have been described and play different roles in smRNA biogenesis and action.

MiRNAs, mainly 21–22 nt in length, are involved in the post-transcriptional regulation of gene expression. Transcribed by the RNA polymerase II from specific genes, the miRNA primary transcripts have an imperfect dsRNA stem-loop structure that is processed by DCL1 into a miR/miR\* duplex. After stabilization



by 3'-O-methylation and transport to the cytoplasm, the mature miRNA is loaded into a RISC complex through the AGO1 or AGO10 proteins (Brodersen et al., 2008). Inside this complex, the miRNA will bind a complementary target RNA, leading to inhibition of its translation or degradation/destabilization (Voïnnet, 2009). In plants, a set of approximately 20 conserved miRNA families, first identified in *Arabidopsis thaliana*, rice or poplar, have been identified in almost all angiosperms studied (Cuperus et al., 2011). Most conserved families are composed of several genes, that code either for a unique mature miRNA or for different but very similar variants. Their targets are also generally conserved among plants (Allen et al., 2004). In contrast, species- or lineage-specific miRNAs are generally present in low amount and encoded by unique genes or small gene families (Cuperus et al., 2011; Turner et al., 2012).

NatsiRNAs arise from natural cis- or trans- antisense overlapping transcripts. One transcript is usually constitutively expressed, while the second is under the control of a promoter responding either to abiotic stresses or pathogen attack (reviewed

in Khraiwesh et al., 2012). Their biogenesis seems quite complex and involved successively DCL2 and DCL1 to produce an original natsiRNA of 22 nt and several secondary 21 nt natsiRNAs (for a more detail description, Zhang et al., 2012).

TasiRNAs are generated from transcripts of non-protein coding genes, called *TAS*. The *TAS* transcripts are first cleaved through the action of specific miRNAs through AGO1 or AGO7. The resulting *TAS* cleavage products are transcribed by RDR6, a RNA dependent RNA polymerase, into long dsRNAs, which are processed by DCL4 into several different 21 nt tasiRNAs following a 21 nt phased interval (Vaucheret, 2005; Voïnnet, 2009; Allen and Howell, 2010). Some tasiRNAs can function in a similar way to miRNAs and regulate genes *in trans* different from their precursors. In *A. thaliana*, to date, four *TAS* genes have been intensively studied (Allen et al., 2005). For instance, *TAS2*-tasiRNAs negatively regulate genes of pentatricopeptide repeat family (PPRs); *TAS3*-tasiRNAs target transcription factors (TF) of the Auxin Response Factor (ARF) family and *TAS4*-tasiRNAs regulate MYB TFs involved in the biosynthesis of anthocyanins.

The last category of siRNA, called hcsiRNAs, are 24 nt in size and mainly derive from heterochromatic DNA regions, transposable elements, regions surrounding centromeres and repetitive sequences. The RNA dependent RNA polymerase RDR2 uses long ssRNA transcribed from heterochromatic regions by the polymerase IV (Onodera et al., 2005), as templates to produce dsRNAs which are processed by DCL3 into several 24 nt hc-siRNAs (Lu et al., 2006). Loaded into a RISC complex containing AGO4, AGO6, or AGO9 proteins, these siRNAs bind complementary DNA loci and induce their methylation through RNA-mediated DNA methylation (RdDM, Havecker et al., 2010; Olmedo-Monfil et al., 2010).

As plant smRNAs play key roles in several developmental stages and in responses to stress, the question of their involvement in the symbiotic nodulation process in legumes has rapidly been addressed by the scientific community. To our knowledge, the first miRNA reported to regulate nodule development (Combier et al., 2006), was miR169, which acts as a negative regulator of *HAP2*. Repression of *HAP2*, which belongs to the CCAAT binding TF family, decreased the number of nodules and altered nodule morphology in the model legume *Medicago truncatula*. Plants that over-expressed one mt-miR169 precursor showed the same phenotype as plants where *HAP2* was inactivated by RNA interference. Combier et al. (2006) proposed that restriction of spatial and temporal expression of the *MtHAP2* target in specific nodule regions was tightly regulated by this miRNA. Afterwards, other miRNAs were associated to nodulation in legumes as reviewed elsewhere (Simon et al., 2009; Khan et al., 2011; Bazin et al., 2012). In parallel, during the five last years, rapid progress in sequencing technologies (from 454 pyrosequencing to SOLEXA and SOLID)

allowed the legume scientific community first, to sequence the genomes and second, to characterize genome-wide a large set of smRNAs from the three legumes: *Medicago truncatula*, *Lotus japonicus* and *Glycine max*. Even though genome sequences of three model legumes are not yet completely finished, a very large portion is available in genomic and EST databases to permit gene identification and annotation. However, we cannot exclude that the lack of certain sequences in these databases may be due to its low expression and presence in non-sequenced regions.

In the first part of this review, we will focus on smRNA diversity in model legumes, present recent data about the miRNA functions in nodulation and define a novel category of siRNAs, the phased siRNAs, likely associated with defence reactions in legumes. In the second part, we mined public genomic databases and published reports to investigate the conservation and specificities of the main components of the smRNA pathways [DICER-like and ARGONAUTE (AGO) proteins] in the three model legumes.

## LEGUME miRNA DIVERSITY AND FUNCTIONS IN NODULATION

### DEEP SEQUENCING REVEALED A LARGE DIVERSITY OF CONSERVED AND NOVEL miRNAs

Since 2006, most miRNAs from animal and plant species have been registered in the miRNA database, called miRBase (www.mirbase.org/, Griffiths-Jones et al., 2006). Taking into account v19.0 (August 2012) for *M. truncatula* and *G. max* and the recently identified miRNAs, from *L. japonicus* (De Luis et al., 2012), we listed 26 conserved miRNA families (Figure 2, Table 1). Most of them (from miR156 to miR399) corresponded to the set

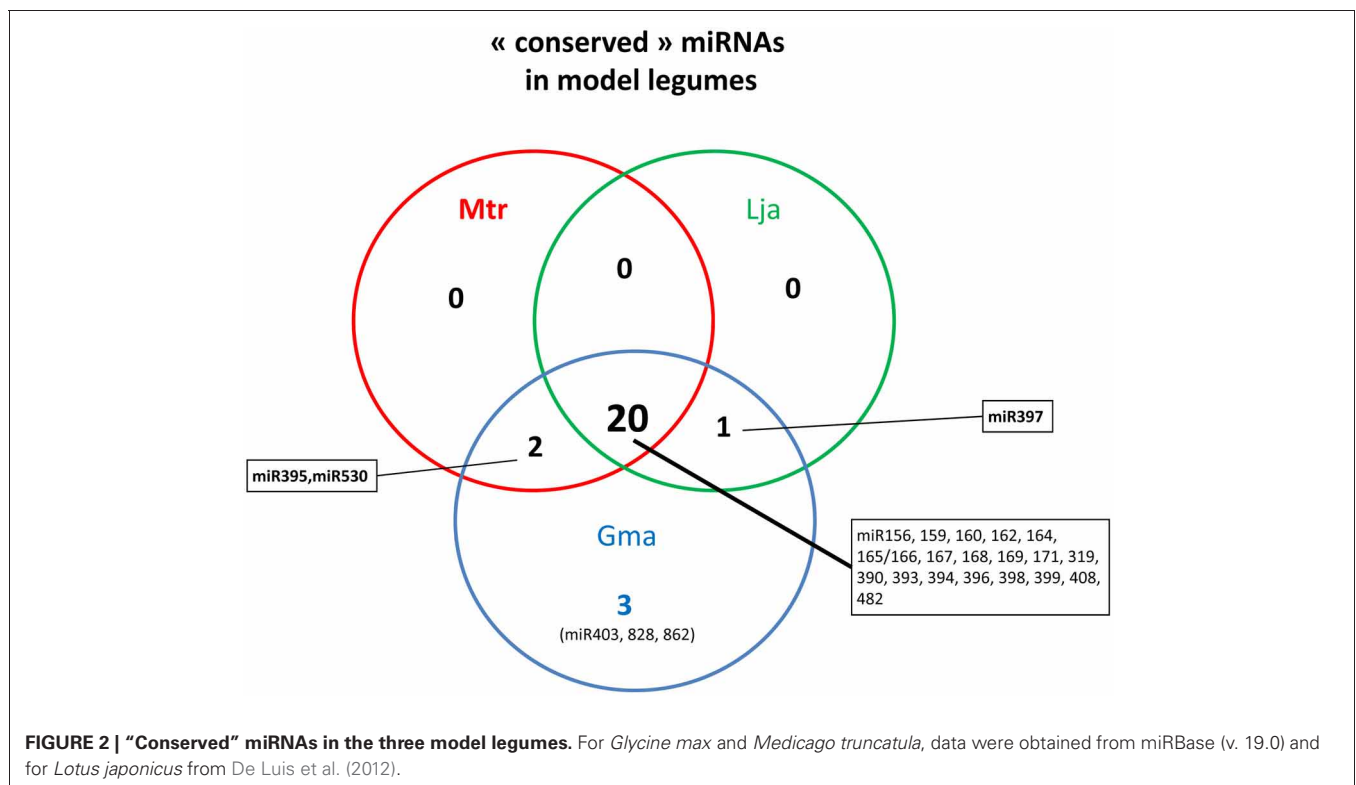


Table 1 | «Conserved» miRNA families in the three model legumes.

miR family	<i>Medicago truncatula</i>		<i>Glycine max</i>		<i>Lotus japonicus</i>	Target cleavage validation	Species number with at least one miRNA gene of the family (miRBase v19)				Functional analyses	
	gene number (miRBase)	smRNA library number	gene number (miRBase)	smRNA library number			eudicots	monocots	conifero-phyta	<i>P. patens</i>		<i>S. moellen-dorffii</i>
156	11		28		P	Mtr,Gma	29	9	1	1	1	N/A
157		1		1	P		5	1				N/A
158		1			P		2					N/A
159	2		6		P	Mtr,Gma	18	10	2		1	N/A
160	6		6		P	Mtr,Gma	24	7	1	1		N/A
162	1		3		P	Mtr	20	4	1			N/A
164	4		11		P	Mtr,Gma	21	6				D'haeseleer et al., 2011
165/166	12		21		P	Mtr,Gma	25	6	3	1	1	Boualem et al., 2008
167	2		10		P	Mtr,Gma	25	7		1		N/A
168	3		2		P	Mtr,Gma	21	6				N/A
169	19		22		P	Mtr,Gma	20	8	1			Combier et al., 2006
170/171	9		21		P	Mtr	25	7	2	1	1	De Luis et al., 2012 ; Ariel et al., 2012 ; Lauressergues et al., 2012
172	3		12		P	Mtr,Gma	25	6				N/A
319	2		14		P	Mtr,Gma	19	5	1	1	1	N/A
390	1		7		P	Mtr	13	4	2	1		N/A
393	2		11		P	Mtr,Gma	15	4				N/A
394		2	7		P	Mtr	16	4				N/A
395	18		13			Mtr	18	5	1	1		N/A
396	3		11		P	Mtr,Gma	25	6	3		1	N/A
397			2		P		15	5	1			De Luis et al., 2012
398	3		3		P	Mtr,Gma	18	6	1			N/A

(Continued)

Table 1 | Continued

miR family	<i>Medicago truncatula</i>		<i>Glycine max</i>		<i>Lotus japonicus</i>	Target cleavage validation	Species number with at least one miRNA gene of the family (miRBase v19)				Functional analyses
	gene number (miRBase)	smRNA library number	gene number (miRBase)	smRNA library number			eudicots	monocots	coniferophyta	P. patens	
<b>399</b>	18		8		P	Mtr	21	6			Branscheid et al., 2011
<b>403</b>			2				12				N/A
<b>408</b>	1		4		P	Mtr, Gma	18	7	1	1	N/A
<b>473-477</b>		1		2	P		7			1	N/A
<b>475</b>				1	P		1				N/A
<b>479</b>					P		5				N/A
<b>482</b>	1		5		P		15	1	3		Li et al., 2010
<b>528</b>					P			5			N/A
<b>530</b>	1		5			Mtr, Gma	10	1			N/A
<b>535</b>					P		6	1	1	1	N/A
<b>828</b>			2				8				N/A
<b>829</b>		1		1	P		2				N/A
<b>837</b>		1			P		2				N/A
<b>841</b>		1			P		2				N/A
<b>857</b>		1			P		3				N/A
<b>858</b>					P		4				N/A
<b>862</b>			2				3				N/A
<b>870</b>		1			P		1				N/A
<b>894</b>		1		1	P					1	N/A

For each conserved miRNA family, the number of genes listed in miRBase v19 in *Medicago truncatula* and soybean (*Glycine max*) are given. If one miRNA was absent from miRBase, we searched it in smRNA libraries or genome reports and indicated the number of reports where it was identified. References of the smRNA sequencing and genome reports used for these additional searches are (Lagadeeswaran et al., 2009; Kulcheski et al., 2011; Lelandais-Briere et al., 2009; Li et al., 2011; Radwan et al., 2011; Szitya et al., 2008; Turner et al., 2012; Wong et al., 2011; Young et al., 2011; Zhai et al., 2011). Presence in *L. japonicus* according De Luis et al. (2012) is indicated by P. Validation of target cleavage by degradome experiments in *M. truncatula* and/or *G. max* were extracted from the following reports: (Devers et al., 2011; Kulcheski et al., 2011; Song et al., 2012; Turner et al., 2012; Zhai et al., 2011; Zhou et al., 2012); N/A, not applicable. P. patens: *Physcomitrella patens* (moss); S. moellendorffii: *Selaginella moellendorffii* (lycophyta) Pale grey boxes: presence of the miRNA, dark gray boxes: miRNA with functional analyses in legumes.



of 21 conserved miRNAs found in nearly all angiosperms (Sunkar and Jagadeeswaran, 2008). However, the absence of 6 conserved families in at least one or even two legumes was unexpected. For instance, miR395 and miR530, two miRNAs regulated by sulfur and nitrogen starvation respectively in *Arabidopsis* (Kawashima et al., 2009; Liang et al., 2010, 2012), were not reported in *L. japonicus*. Only 3 smRNA libraries have been sequenced in this species, and miRNAs having low accumulation levels in those samples may have been missed. We thus searched these miRNAs in available *L. japonicus* genomic data and identified 5 miR395 and one miR530 genes. In contrast, even after deep sequencing of a large variety of tissues and conditions, miR397 was never reported in *M. truncatula*. Lack of this miRNA was first noticed by Sunkar and Jagadeeswaran (2008), who performed an *in silico* prediction of conserved miRNAs in genomic sequences from 682 species, including the three model legumes. In agreement with those observations, no miR397 sequence was found in the last version of *M. truncatula* genome (Young et al., 2011). Putative roles of miR397 in *G. max* and *L. japonicus*, which develop determinate nodules, will be discussed later. Finally, three conserved miRNAs, first described in *A. thaliana*, were only found in *G. max* (miR403, miR828 and miR862) but not in the other legumes. Among them, miR403 is generally considered as dicot-specific and was found in 16 non-legume species. This miRNA may have been lost in certain legume lineages. The presence of miR862 in soybean was striking as, until then, this miRNA was only reported in the *Arabidopsis* genus and in tomato (Gu et al., 2010). In this latter species, miR862 exhibited differential accumulation in roots submitted to phosphate starvation and interacting with arbuscular mycorrhizal (AM) fungi. Hence, although the majority of the conserved miRNAs are present in the three model legumes, specific loss or gains of miRNA genes occurred in certain legume lineages or species.

Conserved miRNAs are generally encoded by multigenic families (Allen et al., 2004). In miRBase, gene number per miRNA family was generally higher in soybean than in *M. truncatula* (Table 1). This was expected as the large soybean genome size, estimated at ~1115 Mbp, was associated to remnants of a whole genome duplication event, which occurred approximately ~13 Mya ago in soybean (Schmutz et al., 2010). Nevertheless, the ratio between gene numbers in *G. max* and *M. truncatula* is generally above two, suggesting that genome duplication was not the only event explaining miRNA diversification in *G. max*. Surprisingly, for some families, an opposite profile was observed: for instance, miR395 and miR399 genes, generally organized in clusters, were more abundant in *M. truncatula* than in soybean (18/13 genes for miR395 and 18/8 genes for miR399 according to miRBase). However, searches of miR395 and miR399-like sequences in the *G. max* genomic database, allowing three mismatches, allowed us to identify 30 and 20 putative members respectively (C. Lelandais, pers. communication), thus suggesting that all genes of these large miRNA families in soybean have not yet been registered in miRBase.

Globally, DCL1 recognize imperfect dsRNA regions like those present in miRNA precursors to release 21–22 nt miRNAs and is the major enzyme involved in miRNA biogenesis. On the other hand, DCL4 also produces 21 nt smRNAs but only from fully

complementary dsRNA. In *A. thaliana*, some miRNAs (miR822, miR839, and miR869) are processed by the action of DCL4 instead of DCL1 (Rajagopalan et al., 2006; Ben Amor et al., 2009). Those miRNAs are present in inverted repeats, share high similarity with their targets and it has been proposed that inverted duplication events formed self-complementary regions which could generate new miRNA genes (Allen et al., 2004). To date, none of these DCL4-dependent miRNAs was described in legumes, reinforcing the hypothesis of their recent and specific origin in the *Arabidopsis* genus. On the other hand, conserved miRNAs can evolve into new miRNA variants that may regulate novel targets. For instance, in *M. truncatula*, a 20 nt variant of miR156 was able to cleave a novel WD40 target in addition to the conserved Squamosa-Binding Protein TF targets (Naya et al., 2010). More recently, a novel isoform of miR171 was discovered in *M. truncatula* and *L. japonicus* (Devers et al., 2011; Bazin et al., 2012; De Luis et al., 2012), that repress a key actor of symbiotic signaling, NSP2 (Nodulation Signaling Pathway 2, a GRAS TF). Further analyses of the miR171 family using both miRBase and comparative genomics (Bazin et al., 2012) revealed that this isoform is also present in non-legume species, such as *Populus trichocarpa* (ptc-miR171) and *Citrus sinensis* (csi-miR171b). However several plants unable to form root endomycorrhiza like *A. thaliana*, *Brassica napus* and the gymnosperm *Pinus taeda*, do not contain this isoform suggesting a role in root symbioses.

Several non-conserved miRNAs, first described in one legume species (soybean: Subramanian et al., 2008; Wang et al., 2009 or *M. truncatula*: Szittyta et al., 2008 and Jagadeeswaran et al., 2009) have been reported as “legume-specific”. In Table 2 we summarize some characteristics of 15 of them selected either because they are present in at least two of the legumes analyzed or because they have been functionally related to nodulation (miR1512, miR1515; miR1521; Li et al., 2010). According to miRBase, all, except miR1511, miR1515, miR2111 and miR2118, may be considered as “legume-specific”. However, Zhai et al. (2011) showed that variants of miR1507 and miR1509 were present in smRNA libraries of non-leguminous species. For instance, miR1507 was highly abundant in grapes. In addition, miR2118 was sequenced in 34 non-legume species, including 4 gymnosperms, suggesting a very ancestral origin (>250 million years). During the last three years, several reports of smRNA library sequencing and genome wide miRNA identification (see Data Sheet 1 for the list) allowed to found hundreds of novel miRNA families in legumes. In miRBase, 229 and 179 novel miRNA families have been registered for *M. truncatula* and *G. max*, respectively, whereas De Luis et al. (2012) reported 35 novel miRNA families in *L. japonicus*. In miRBase, *M. truncatula* and *G. max* correspond to the first and third plant species in term of total numbers of miRNA genes, with 675 and 506 genes, respectively. This huge diversity, comparable to, e.g., rice, may be due to the high number of smRNA deep sequencing analyses performed (9 and 10 reports for *M. truncatula* and *G. max*, respectively, listed in Data Sheet 1). In addition, likely “false” candidates have been registered due to low stringency criteria used for miRNA identification in some of the initial studies. For instance, between 2009 and 2011, many miRNAs (including certain identified in our

Table 2 | Selection of 15 “legume” miRNA families.

family	<i>Medicago truncatula</i>		<i>Glycine max</i>		<i>Lj</i>	other legumes (miRBase)	non legumes (miRBase)	miR* reported	target validation D: degradome R: RACE-PCR N/A: not applicable
	gene number (miRBase)	number of sRNA reports	gene number (miRBase)	number of sRNA reports					
1507	1	7	3	2		<i>Gso, Vvu</i>		<i>Mt, Gm</i>	D
1508			3	2		<i>Gso</i>		<i>Gm</i>	N/A
1509	2	6	2	2	P	<i>Gso</i>		<i>Mt, Gm</i>	D
1510	2	5	2	2	P	<i>Gso</i>		<i>Mt, Gm</i>	D
1511		1	1	2	P		<i>Mdo</i>	<i>Gm</i>	D
1512			3	2				<i>nd</i>	N/A
1515			2	2			<i>Csi</i>	<i>nd</i>	R
1520	0	1	18	1				<i>Mt</i>	N/A
1521			2	1				<i>nd</i>	N/A
2086	1	7				<i>Aau</i>		<i>Mt</i>	D
2109		2	1	2	P	<i>Gso</i>		<i>Gm</i>	D
2111	21	6	7	2	P	<i>Lja,</i>	<i>Ptc, Ath, Vvi, Mdo</i>	<i>Mt, Gm</i>	D
2118	1	5	2	3		<i>Gso, Pvu, Vvu</i>	<i>Mdo, Osa, Zma</i>	<i>Mt</i>	D
2119	1	5	1	1	P	<i>Pvu</i>		<i>Mt</i>	D
2199	1	3			P	<i>Pvu</i>			N/A

For selected “legume” miRNA families, the number of genes listed in miRBase v19 in *Medicago truncatula* and soybean (*Glycine max*) species are given. The number of smRNA sequencing or genome reports where it was cited is also indicated. Reports of smRNA sequencing and genome sequencing used for these additional searches are: (Jagadeeswaran et al., 2009; Kulcheski et al., 2011; Lelandais-Briere et al., 2009; Li et al., 2011; Radwan et al., 2011; Szittyta et al., 2008; Turner et al., 2012; Wong et al., 2011; Young et al., 2011; Zhai et al., 2011). Presence in *L. japonicus* (*Lj*) according to De Luis et al. (2012) is indicated by P. Validation of target cleavage by degradome experiments in *M. truncatula* and/or *G. max* were extracted from the following reports (Devers et al., 2011; Kulcheski et al., 2011; Song et al., 2011; Turner et al., 2012; Zhai et al., 2011; Zhou et al., 2012). *P. patens*: *Physcomitrella patens* (moss); *S. moellendorffii*: *Selaginella moellendorffii* (lycophyta). Other Legumes species: *Pvu* (*Phaseolus vulgaris*); *Vvu* (*Vigna unguiculata*); *Ahy* (*Arachis hypogea*); *Aau* (*Acacia auriculiformis*); *Gso* (*Glycine soja*); Non legume species cited in the table: *Ath* (*Arabidopsis thaliana*); *Mdo* (*Malus domestica*); *Osa* (*Oryza sativa*); *Ptc* (*Populus trichocarpa*); *Vvi* (*Vitis vinifera*); *Zma* (*Zea mays*). Pale grey lines: legume-specific miRNAs (according miRBase data); dark grey boxes: miRNA with functional analyses in legumes (Li et al., 2010).

work, Lelandais-Briere et al., 2009), were registered in miRBase, although no miR\* was sequenced in the libraries, a criteria that became required for bona fide miRNAs (Meyers et al., 2008).

In this context, a major challenge in the next years will be to select the “best” novel legume miRNAs for functional analyses. Although several nodulation-responsive miRNAs have already been identified (Subramanian et al., 2008; Simon et al., 2009; Li et al., 2010; De Luis et al., 2012; Turner et al., 2012), microarray based genome-wide transcriptional analyses as well as robust statistical comparisons of smRNA abundancies in libraries from various developmental stages or tissues will be very helpful. In addition to miRNA expression, it is of interest to analyse their target mRNAs. Recent analyses used “degradome” experiments (also called PARE, German et al., 2009) to detect enrichment of cleaved mRNAs on miRNA complementary sites. By analysing changes in miRNA expression patterns which correlate with specific degradation of mRNA targets, we can assess the potential regulation by miRNAs of several mRNA targets, both in *G. max* (Song et al., 2011; Turner et al., 2012) and *M. truncatula* (Devers et al., 2011; Zhai et al., 2011; Zhou et al., 2012). For instance, the mRNAs encoding a salt-tolerance protein and a xyloglucan endo-transglucosylase/hydrolase (an enzyme that participates in cell wall formation) are targeted by

miR2708 and miR2687, respectively, in response to the presence of mercury (Zhou et al., 2012). In addition, mtr-miR2681 targets several transcripts coding TIR-NBS-LRR resistance proteins and cleavage of five of them (TC127116, TC115294, TC128879, BG587250, and NP7251801) were identified in roots treated with 10  $\mu$ M Hg (Zhou et al., 2012). Degradome data in mycorrhizal roots or control conditions showed that genes associated to defense responses (Medtr6g098880.1, Medtr2g046350.1, and TC138295) are specifically regulated by miR5213 or miR2678, respectively (Devers et al., 2011). Interestingly, miR5213 was conserved among species able to interact with AM fungi (Devers et al., 2011). Finally, this degradome allowed Devers et al. (2011) to propose that some miR\* (complementary smRNA to the miRNA generated by DCL processing) accumulating at high levels in *M. truncatula*, were able to cleave complementary mRNA targets. The miR169\* cleaved *MtBCPI* transcripts coding for an arbuscule-specific protein in mycorrhizal roots whereas cleavage products of a GRAS TF predicted as a target of miR5204\* were also sequenced in mycorrhizal roots (Devers et al., 2011).

#### RECENT LEGUME miRNAs LINKED TO NODULATION

Differential accumulation of miRNAs during rhizobial interactions and nodule development have been reported in several

studies (Subramanian et al., 2008; Wang et al., 2009; Li et al., 2010). However, functional analyses of miRNAs remained rare and only two miRNAs, miR169, and miR166, were experimentally associated to nodule development before 2009 (Data Sheet 2; Comber et al., 2006; Boualem et al., 2008). During the last three years, a set of interesting studies have interconnected additional miRNAs to the regulatory networks that control nodulation (Data Sheet 2). D'Haeseleer et al. (2011) reported that over-expression of miR164, a conserved miRNA targeting *NAC1* TF in roots, affected nodule organogenesis in *M. truncatula*, presumably through deregulation of auxin responses. A functional analysis of several soybean miRNAs highly expressed in roots inoculated by symbiotic bacteria, first identified by Subramanian et al. (2008), was performed by Li et al. (2010): the conserved miR482 and five legume-specific miR1507, miR1511, miR1512, miR1515, and miR1521. Analysis of miRNA accumulation in the NOD49 nodulating mutants (mutated in the NOD factor receptor 1, *NFR1*) and the super-nodulation mutant NTS382 (impaired in *NARK1*, a receptor kinase involved in nodule auto-regulation pathway), pointed out that *gma-miR1507*, *gma-miR1511*, and *gma-miR1512* expression was dependent on *NFR1* or *NARK1*. Furthermore, miRNA over-expression in transgenic roots under the control of a constitutive or a nodulation-specific *ENOD40* promoters, showed that miR482, miR1512, and miR1515 positively regulated nodule number (Li et al., 2010). MiR1512 targets a Copine-like membrane protein that participates in cell signaling and transport. MiR482 and miR1515 repress disease resistance genes and/or *DCL2*, the DICER-like gene mainly related to defence against viruses. This could be related to the fact that symbioses share some common features with early defence responses (Ochman and Moran, 2001; Simon et al., 2009; Corradi and Bonfante, 2012; Marchetti et al., 2010; Bourcy et al., 2013; Peleg-Grossman et al., 2013). In *G. max*, 4 novel miRNAs (*gma-new-miR4416a*, *gma-new-miR4416b*, *gma-new-mi13587*, and *gma-new-miR50841*) were also reported as highly expressed in nodules (Turner et al., 2012). In addition, the accumulation of the corresponding targets in roots and nodules negatively correlated with the relative abundance of the miRNAs, suggesting that these genes may be specifically regulated in organs through miRNA spatial distribution (Turner et al., 2012). However, their functions remain unknown.

Very recently, the first identification of miRNAs from *Lotus japonicus* has been published by De Luis et al. (2012). These authors showed that miR397 is required for the establishment and maintenance of determinate nodules in *Lotus* but not for nodule organogenesis. The *snf* mutant (spontaneous nodulation mutant) inoculated with *Mesorhizobium loti* accumulated more miR397 than non-inoculated plants. One potential miR397 target in *M. truncatula* is homologous to the *A. thaliana* LACCASE10 (a copper-containing oxidase enzyme). In *A. thaliana*, miR397 was linked to nutrient interchanges between shoots and roots. During nodulation, nutrient exchanges occur between plant cells and bacteroids in nodules, and miR397 may be necessary to maintain a correct level of copper during this process (De Luis et al., 2012). Moreover, to maintain the nitrogen fixation rate, O<sub>2</sub> concentrations must be regulated inside the nodule cells. Cu/Zn superoxide dismutase (SOD) scavenges superoxide

radicals avoiding the inhibition of nitrogenase activity (Rubio et al., 2007). This enzyme depends on the availability of Cu<sup>2+</sup> inside the cells and a decrease in Cu/ZnSOD expression level takes place when miR397 is over-expressed. This result is consistent with the fact that miR397 levels are higher in mature-senescent nodules than in younger ones, a stage where over-production of reactive oxygen species can be detected (Matamoros et al., 1999). Interestingly, miR397 expression is regulated by SPL7 in *Arabidopsis*. This TF also controls the expression of miR398, another conserved miRNA which negatively regulates Cu/ZnSOD genes (Mendoza-Soto et al., 2012). In addition, miR408 and miR857 were also regulated in response to Cu<sup>2+</sup> through SPL7; although their level does not increase during nodulation (De Luis et al., 2012). In conclusion, miR397 and miR398 may participate to a complex regulatory mechanism that controls at least copper homeostasis in nodules. As indicated before, although miR397 has been reported in soybean and *L. japonicus* (Wang et al., 2009; De Luis et al., 2012) there is no evidence of this miRNA in *M. truncatula*. Thus, miR397 could be necessary for bacterial infection in determinate nodules like those from *L. japonicas* and *G. max* but not in the indeterminate nodules of *M. truncatula*. There are several differences between these nodule types including the persistence of the meristem and bacteroid differentiation and several genes were specifically associated to indeterminate nodules (Van de Velde et al., 2010). Differences in miRNAs between determinate and indeterminate nodules are consistent with the fact that both types of legumes do not respond equally to nodulation signals and show different morphology, physiology and responses to stress (Subramanian et al., 2007; Deinum et al., 2012; López-Gómez et al., 2012).

Finally, the role of miR171 in legumes, first suggested by Devers et al. (2011), emerged through three different reports. Laouressgues et al. (2012) and De Luis et al. (2012) showed that the miR171h and miR171c variants are fundamental to establish symbiotic mycorrhization and nodulation in *M. truncatula* and *L. japonicus*, respectively. When *M. truncatula* roots are infected by the AM fungus *Rhizophagus irregularis*, the increase in miR171h was followed by a concomitant decrease in the corresponding target *NSP2*. In addition, overexpression of miR171h led to a decrease in fungal colonization associated to the down-regulation of mycorrhizal marker genes (Laouressgues et al., 2012). In *L. japonicus*, De Luis et al. (2012) showed that, similarly to miR397, the miR171c isoform accumulated in inoculated *snf* mutants and that this miRNA variant was associated to nodule establishment and maintenance but not organogenesis. In both species, the specific miR171 isoforms studied were able to cleave the target transcript *NSP2*, a TF involved in NOD factor signaling. Furthermore, as *MtNSP2*, *mt-miR171h* was also early activated in response to cytokinins (Ariel et al., 2012). The emerging idea is that *NSP2* evolved in legumes to acquire specific functions during nodulation. Therefore, miR171h may be required for nodule but also for mycorrhiza establishment through *NSP2* regulation, a node involving cytokinin signaling. The common function of miR171 isoforms in nodulation and mycorrhization reinforces the idea that nodule development may have evolved from AM



fungi interactions through a diversification and specialization process.

### PHASED siRNA IN LEGUMES

Several siRNAs were identified showing “phasing” or processing cleavage of 21 nt derived from a long precursor. This is the case for the *TAS* genes that are targeted by different miRNAs as well as for the recently described phasiRNAs.

### TAS3 tasiRNA REGULATION IS CONSERVED IN LEGUMES

In *Arabidopsis thaliana*, 4 *TAS* genes have been well described (*TAS1* to *TAS4*). *TAS1* is target of miR173, while *TAS2* and *TAS4* are cleaved by miR828. Their miRNA-dependent cleavage requires the action of AGO1 (Peragine et al., 2004). In contrast, *TAS3*-tasiRNA biogenesis depends on miR390 and AGO7 (Montgomery et al., 2008). The cleaved *TAS* transcripts are used as templates by RDR6 in complex with the Suppressor of Gene Silencing III protein (SGS3). The resulting long dsRNA is cleaved into secondary 21 nt siRNAs respecting a certain phase due to the processivity of DCL4 (Vaucheret, 2005). Hence, DCL4 generates several phased tasiRNA that will be loaded into an AGO1-containing RISC complex to target complementary mRNAs. Interestingly, except for miR390, the miRNAs that participate in the biogenesis of tasiRNA are 22 nt in length (Cuperus et al., 2011). This 1 nt difference of the loaded miRNA has been proposed to be the clue to direct AGO1 from mRNA cleavage without any amplification (like all 21 nt miRNAs) to the production of secondary siRNAs. Recently, Manavella et al. (2012) proposed that the asymmetry provoked by the loading of a 22 nt miRNA would alter AGO1 conformation and, in this way, it might condition the capacity to trigger transitivity via RDR enzymes.

The most conserved *TAS* gene in plants is *TAS3* which is targeted by miR390, a 21 nt miRNA. *TAS3* RNA presents two miR390 binding sites, however only the 3' site is cleaved following base pairing with miR390 (Montgomery et al., 2008). Furthermore, miR390 is loaded on a specific AGO7-containing complex to trigger tasiRNA production on cleaved *TAS3* transcripts. *TAS3*-derived tasiRNAs target specific members of the ARF family: ARF2, ARF3 and ARF4. These TFs are involved in auxin responses and therefore in many developmental stages (Jouannet et al., 2012), notably lateral root formation a process which may be linked to symbiotic interactions in legume roots (Marin et al., 2010). In *M. truncatula*, Jagadeeswaran et al. (2009) reported that, among the four *Arabidopsis* *TAS* genes, only *TAS3* and its *ARF* targets are conserved. In addition, mutants affected in AGO7 and SGS3, two main enzymes involved in *TAS3*-tasiRNA biogenesis were found in *L. japonicus*. These mutants, called *Ljrel* (REduced Leaflet), display similar leaf phenotypes, with abaxialized leaflets and lower leaflet numbers as well as defects in flower development (Peragine et al., 2004; Yan et al., 2010). *Ljrel1* mutants (affected in SGS3 function) down-accumulate *TAS3*-tasiRNAs leading to the misregulation of their *ARF* targets (Yan et al., 2010). However, although *TAS3*-siRNA are involved in the quantitative regulation of root lateral organogenesis in *A. thaliana* (Marin et al., 2010), nodulation phenotypes were not yet reported in these mutants.

### THE PHASED SHORT INTERFERING RNA (phasiRNA): A NEW IMPORTANT CLASS OF siRNA IN LEGUMES

As mentioned before, 22 nt miRNAs (instead of their 21 nt corresponding variants) may affect the conformation of AGO1 and trigger the production of secondary phased siRNA of 21 nt (Chen et al., 2010; Cuperus et al., 2011). In *Arabidopsis*, similar abundances of the mature 21 nt and 22 nt variants were found for several miRNA families, like miR173, miR828, miR472 (Cuperus et al., 2011) or miR319 and miR771 (Chen et al., 2010). The 22 nt variants are loaded into an AGO1-containing RISC complex and triggers cleavage of their complementary targets. However, like for *TAS* transcripts, cleavage products are transformed into long dsRNA by the RDR6/SGS3 complex and then spliced into phased secondary siRNAs, mainly through DCL4 (Cuperus et al., 2011). Such a mechanism has also been shown in tobacco (Li et al., 2012). In rice, biogenesis of phased siRNAs occurred through a similar mechanism, which involves 24 nt siRNAs and a DCL3 isoform instead of DCL4 (Song et al., 2012). Indeed in the *osdcl4* mutant, 24 nt phased smRNA still accumulate thanks the action of the OsDCL3b protein. In that species, two 22 nt miRNAs (miR2118 and miR2275) were necessary to induce the secondary production of 24 nt phased siRNAs. Hence, at least in rice, the production of phased siRNAs depends on a particular DCL isoform, different from DCL4 and issued from DCL3 duplication and specialization (Song et al., 2012). Those results indicate that 22 nt miRNAs may have been selected to generate both 21 or 24 phased siRNA in plants during evolution.

Last year, Zhai et al. (2011) studied in detail the diversity of 21 nt phased siRNAs in several smRNA libraries from soybean and *M. truncatula*. These siRNAs, called phasiRNA, derive from *PHAS* genes, that are primarily cleaved by 22 nt mature miRNAs. These authors identified 114 and 41 *PHAS* loci in *M. truncatula* and soybean, respectively. In *M. truncatula*, 112 *PHAS* loci corresponded to protein-coding genes and 2 to intergenic regions, while 26 and 15 were identified as protein-coding genes or intergenic respectively in soybean. Around 68% of the *PHAS* loci in *M. truncatula* contained one 22 nt miRNA binding site, and most of them were triggered by one of the four 22 nt miRNAs, miR1501, miR1509, miR2109, and miR2118, which are predominantly abundant in this species. Among *PHAS* loci which contained two 22 nt miRNA binding sites, those authors found an *APETALA2* (*AP2*)-like gene, which possess one cleavage site for miR172 and a predicted non-cleavable miR156 target site, resembling the two miR390 complementary motives in *TAS3*. In soybean and *L. japonicus*, however, *AP2* orthologous only present one miR172 binding site. Thus, the acquisition of the second (miR156) miRNA binding site in *MtAP2* certainly happened recently in evolution (Zhai et al., 2011) and this may have consequences on the generation of secondary siRNAs.

Interestingly, some genes coding for enzymes involved in smRNA pathways have been identified as *PHAS* loci (Zhai et al., 2011). In soybean, *GmSGS3a* transcripts are targets of miR2118. *DCL2* mRNAs are cleaved by miR1507 or miR1515 in *M. truncatula* and soybean, respectively. This suggests that in those legumes, *DCL2* genes evolved independently to acquire similar regulation by a 22 nt miRNA and production of secondary phasiRNAs. The targeting activity of the phasiRNAs on genes involved in smRNA

biogenesis may also suggest a feedback mechanism, reminiscent of the regulation of *AGO1* and *DCL1* by miR168 and miR162, respectively. Hence, 22 nt miRNAs and phasiRNAs generated by *DCL2* or *SGS3* may control these targets which are essential for their own production. Jagadeeswaran et al., 2009 identified TIR-NBS-LRR proteins as mt-miR2118 targets. In addition, recently, it has been shown that most *PHAS* loci targeted by miR1507 and miR2118 encode NBS-LRR resistance proteins (Zhai et al., 2011) leading the authors to suggest an important role of these miRNAs in the control of biotic interactions in legumes, including symbiotic nodulation. However, these miRNAs are not specific from legumes and similar regulations of NBS-LRR genes by 22 nt miRNAs were described in non-legumes, like *Nicotiana benthamiana* and tomato (Li et al., 2012). In the first species, ten 22 nt miRNA families target a large set of resistance genes. Among them, Nta-miR6019, also found in other solanaceous species, conferred resistance to the TMV (Tobacco Mosaic Virus). In tomato, the 22 nt variant of miR482 cleaves six NBS-LRR proteins, and at least two of them contain a second miR482 binding site (Shivaprasad et al., 2012). Infection of tomato plants by Turnip crinkle virus (TCV), Cucumber mosaic virus (CMV) or Tobacco rattle virus (TRV) increased the expression of NBS-LRR by decreasing miR482 accumulation (Shivaprasad et al., 2012). The authors proposed that, when the pathogen is absent, phasiRNAs block the NBS-LRR expression in order to reduce energetic cost to the plant. These data suggested that the control of disease resistance gene expression by 22 nt miRNAs through secondary production of phasiRNAs evolved early in plants and is not exclusive from legumes. As in nature, bacteria and virus pathogens have co-evolved with plants and gene-for-gene relationship exist (Jones and Dangl, 2006), virus infection could trigger a suppression of NBS-LRR-targeting miRNA biogenesis which leads to increase NBS-LRR expression and defense reactions. Symbiotic interactions share some common features with pathogenesis (Soto et al., 2009) and co-evolution may have favored the maintenance of miRNA targeting of NBS-LRR phased loci in symbiotic interactions.

In addition to the function of new miRNAs evolved in legumes and regulated in symbiotic interactions, functional analyses of those 22 nt miRNAs and their derived phasiRNAs in nodulation and pathogen responses remain an interesting challenge for the next years. Furthermore, the 24 nt heterochromatic siRNAs generally associated to the establishment of epigenetic patterns in different processes have been largely remain a forgotten aspect of smRNAome diversity in legumes.

### SMALL RNA PATHWAYS IN LEGUMES

Despite the diversity of smRNA in different plants, the core of their biogenesis pathways depends generally on DCL and AGO proteins (Parent et al., 2012). DCLs and AGOs are encoded by multigenic families of 10 and 4 members in *A. thaliana* and 19 and 6 members in rice (Kapoor et al., 2008). In comparison, vertebrates, nematodes and yeast contain only one DCL, whereas insects, protozoa and fungi have two DCL isoforms. Plants thus globally show an important gene diversification of DCLs and AGOs, that led to the specialization of certain isoforms in different smRNA pathways (Hutvagner and Simard,

2008; Murphy et al., 2008; Liu et al., 2009). Although previous works have been published about phylogenetic and evolutionary aspects of these enzymes in plants, to our knowledge, there is no published global analysis of DCL and AGO proteins focused on model legumes. Even though their genome sequences are not yet completely finished, a very large portion is available in genomic and EST databases for database mining. For this, we carried out TBLASTX analyses to identify the orthologs of *Arabidopsis* DCL and AGO proteins in the three legumes for which large genomic databases are available: *Lotus japonicus* (Miyakogusa.jp 2.5, Sato et al., 2008), *Medicago truncatula* (Mt 3.5.1, Young et al., 2011), and *Glycine max* (Glyma1.181, <http://www.plantgdb.org/GmGDB/>). In addition, we compared our results to those already published for rice (Kapoor et al., 2008) and poplar (*Populus trichocarpa*, Margis et al., 2006). The complete list of DCL and AGO genes is given in Data Sheet 3.

### DICER-LIKE PROTEINS IN LEGUMES

In *A. thaliana*, the four DICER-like proteins produce differently sized smRNAs and have complex relationships (Gascioli et al., 2005). DCL1 is mainly involved in miRNA biogenesis. It processes the miR/miR\* duplexes from imperfect fold back stem-loops of the pri-miRNA precursors, while DCL2, DCL3 and DCL4 are more generally responsible for generating siRNA from dsRNA originating from exogenous elements, natural antisense genes, TAS (and likely PHAS) transcripts or repeated heterochromatic regions. *dcl2/3/4* triple mutants showed a reduction in siRNA production, but there was no change in miRNA populations, confirming that DCL1 is the main enzyme involved in miRNA biogenesis. In addition, in rice and *A. thaliana*, DCL1 loss of function causes embryo lethality (Liu et al., 2005; Song et al., 2011). However, in *A. thaliana*, smRNAs derived from inverted repeats that fold into imperfect hairpins were also more abundant in *dcl2/3/4* triple mutant (Henderson et al., 2006). This observation correlates with the fact that DCL1 is necessary for the accumulation of the Cauliflower mosaic virus-derived siRNA that also derive from an imperfect fold back RNA structure (Dunoyer et al., 2007). Thus, DCL1 is specialized in the production of miRNA or siRNA of 21 nt from imperfect fold backs structures (Chapman and Carrington, 2007). In *A. thaliana*, Gascioli et al. (2005) showed that the accumulation of tasiRNA targets increased in *dcl4* and *dcl3/dcl4* mutants, while *dcl2/dcl3* and *dcl3/dcl4* present stochastic developmental phenotypes due to the lack of accumulation of heterochromatic siRNA-directed marks. On the other hand, analysis of *dcl* double and triple mutants pointed out that DCL2, DCL3 and DCL4 have compensatory functions among them (Gascioli et al., 2005; Henderson et al., 2006). For instance, in *dcl3* mutants, DCL2 and DCL4 are able to produce 22–21 nt siRNAs from DCL3 substrates (Gascioli et al., 2005). In addition, in viral siRNA biogenesis, DCL4 acts in a hierarchical manner with DCL2 (Rajagopalan et al., 2006), using the DCL2-dependent 22 nt siRNA to trigger secondary siRNA biogenesis (Chen et al., 2010).

Other plants than *A. thaliana* generally possess more than 4 DCLs. For instance poplar (*P. trichocarpa*) has 5 DCLs (PtDCL1, PtDCL2a, PtDCL2b, PtDCL3, and PtDCL4, Margis et al., 2006) and rice contains 6 DCLs (OsDCL1, OsDCL2a, OsDCL2b,

OsDCL3a, and OsDCL3b and OsDCL4) (Kapoor et al., 2008). Although Kapoor et al. (2008) reported 3 different *OsDCL1*-like genes, we assume that the so-called *OsDCL1b* and *OsDCL1c* genes in fact encode another kind of ribonuclease III, closed to DCLs, the RNase Three Like (RTL) proteins. Thus, only the *OsDCL1a* gene was taken into account in our study. As previously reported, in rice, duplication of *DCL3* has been followed by a specialization of the two isoforms. Indeed, OsDCL3a classically triggers hc-siRNA biogenesis while OsDCL3b, which shows a specific expression in early stages of seed development (Kapoor et al., 2008), is mainly involved in the biogenesis of 24 nt phased siRNA (Song et al., 2012). Accordingly to this, inactivation of *OsDCL3b* by RNA interference affected the accumulation of 24 nt phased siRNAs but not of 24 nt hc-siRNAs (Song et al., 2012).

In legumes, Curtin et al. (2012) reported 6 DCL genes in *G. max* (*GmDCL1a*, *GmDCL1b*, *GmDCL2a*, *GmDCL2b*, *GmDCL4a*, and *GmDCL4b*). In *M. truncatula*, both Capitão et al. (2011) and Young et al. (2011) identified one homolog for *DCL1*, *DCL2* and *DCL3* genes, but pointed out the absence of *DCL4* homolog. By searching in *L. japonicus* genomic sequences, we identified five *LjDCL* genes: *LjDCL1*, *LjDCL2a* and *LjDCL2b*, *LjDCL3* and *LjDCL4* (Data Sheet 3). To construct a phylogenetic tree of DCLs (Figure 3), only the complete predicted proteins were retained. According to our analysis, DCL proteins were clearly grouped into four monophyletic groups, each corresponding to one of the four AtDCLs, and with high sequence conservation between species. Like in non-legumes, only one *DCL1* gene was found in *M. truncatula* and *L. japonicus*, while two *GmDCL1* genes are present in soybean. The functional significance (if any) of this specific event of *DCL1* duplication in soybean remains to be elucidated. In contrast, like in many angiosperms, a duplication of *DCL2* was observed both in *L. japonicus* and *G. max*. According to Margis et al. (2006), this duplication event took place before the divergence between *P. trichocarpa* and *A. thaliana*. However, *M. truncatula* apparently contains a unique *DCL2* gene. As the genome sequence is not fully complete in that species, the most probable explanation may be that the second *DCL2* locus has not yet been sequenced. However, we were also not able to identify expressed sequences corresponding to a putative second *DCL2* homolog, although the *M. truncatula* EST database is very large (TIGR MtGi11.0). In contrast to rice and maize, that shared a *DCL3* duplication (Margis et al., 2006; Kapoor et al., 2008), only one *DCL3*-like gene was described in the three model legumes, even in the *G. max* paleopolyploid genome (Gill et al., 2009). Curtin et al. (2012) reported a second *DCL3* but the presence of an in-frame stop codon and the absence of several DCL characteristic domains led to consider it as a pseudogene. These data thus reinforces the idea that *DCL3* diversification may be specific to monocots.

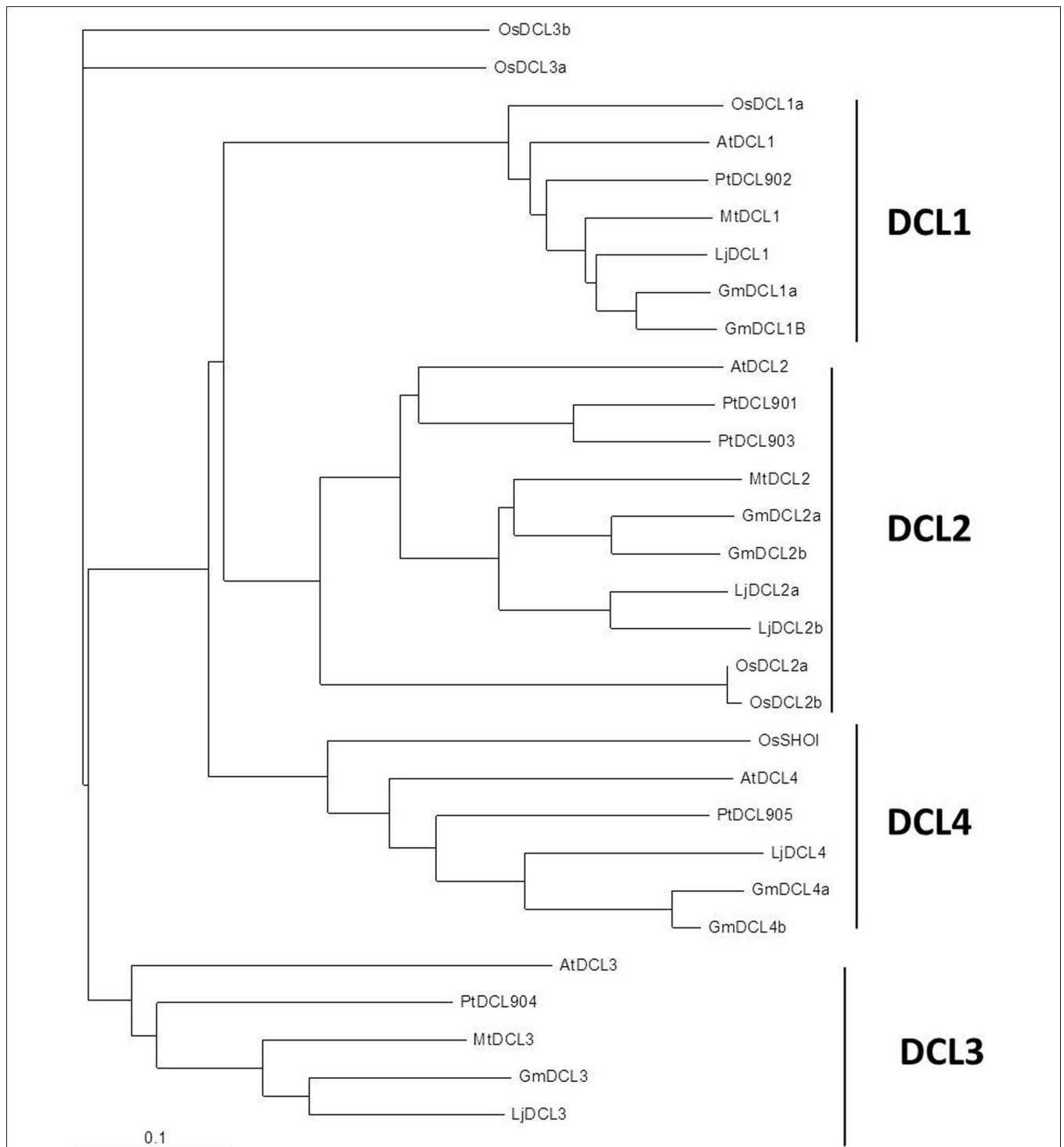
To our opinion, the absence of *DCL4* sequences in the *M. truncatula* genome annotation (Young et al., 2011) is certainly due to its still incompleteness. Indeed, we were able to found a genomic contig (contig\_162690\_1.1) which contains a putative incomplete *DCL4* gene. Unexpectedly, two *DCL4* homologs were found in soybean, that shared a similarity of 79.9% and 80.3% at nt and amino acid levels, respectively. The genome of soybean has suffered two rounds of large scale-sequence or segmental duplication

(Gill et al., 2009), thus the existence of two *DCL1* and two *DCL4* genes in soybean may derive from a segmental duplication event. This idea is reinforced by genome browser data, showing that both *DCL4s* are flanked by genes coding homologous proteins (at 5', Glyma13g22420 and Glyma17g11220, which are annotated as MCM5 minichromosome maintenance family protein; and at 3', Glyma13g22460 and Glyma17g11250, annotated as Phosphoinositide binding protein).

To further investigate the diversity of legume DCLs, we compared their conserved domains using Simple Modular Architecture Research Tool-SMART version 7 (Letunic et al., 2012; Figure 4). In plants, DCLs contain six types of conserved domains: helicase-C, DEAD-helicase box (DEXD/H-box), DUF283, RNase III (RIBOc), double-strand RNA-binding (dsRBD) and PAZ (Piwi/Argonaute/Zwille) domains (Margis et al., 2006; Murphy et al., 2008). Two RNase III domains are required to constitute an intramolecular dimer, necessary for the cleavage of dsRNA substrates. According the DCL, one or two dsRBD domains are present. The PAZ domain participates in protein-protein interactions and, because of its ubiquitous presence, a role in the interaction between DCLs and other proteins was suggested, thus guiding template recognition (Carmell and Hannon, 2004). The DEAD-helicase box domain seems to be essential in DICER auto-regulation. In humans, removal of this domain increases the cleavage rate of DICER proteins. DUF283 domain may be involved in the selection of the different smRNA biogenesis pathways as it recognizes the asymmetry of dsRNA substrates (Liu et al., 2009).

In the plants studied so far, *DCL1* and *DCL4* proteins possess all functional domains, including two dsRBD domains. As expected, in legumes, *DCL1s* were also highly conserved and shared all domains necessary for miRNA biogenesis (Gascioli et al., 2005; Parent et al., 2012). Unexpectedly, only one dsRBD domain was found in the two *GmDCL4* isoforms. However, a *DCL4*-like gene coding for a protein containing one dsRBD domain is located 2654 bases from the presumed stop codon of *GmDCL4a*, suggesting a problem of annotation. In contrast to *DCL1* and *DCL4*, *DCL2* and *DCL3* isoforms are much more variable, in particular in their number of dsRBD domains (Cerutti and Casas-Mollano, 2006; Margis et al., 2006). In rice, for instance, both *DCL2* isoforms, although functional, contain only one dsRBD domain (Margis et al., 2006). In addition, OsDCL3a has two classic dsRBD domains, whilst OsDCL3b lacks one. Although Margis et al. (2006) reported that AtDCL3, PtDCL3, and OsDCL3b possess two dsRBD domains, according to our analyses using the SMART version 7 software (<http://smart.embl-heidelberg.de/>), those proteins appear to contain only one typical dsRBD domain. In legumes, we also noticed such variability with two dsRBD domains in MtDCL3 and only one in *GmDCL3* and *LjDCL3*. In our opinion, the more striking observation in term of domain composition was the lack of any canonical dsRBD in *LjDCL2a*. Although a problem of annotation remains an hypothesis, it is possible that the *DCL2* function is mainly assumed by *LjDCL2b* in *L. japonicus*.

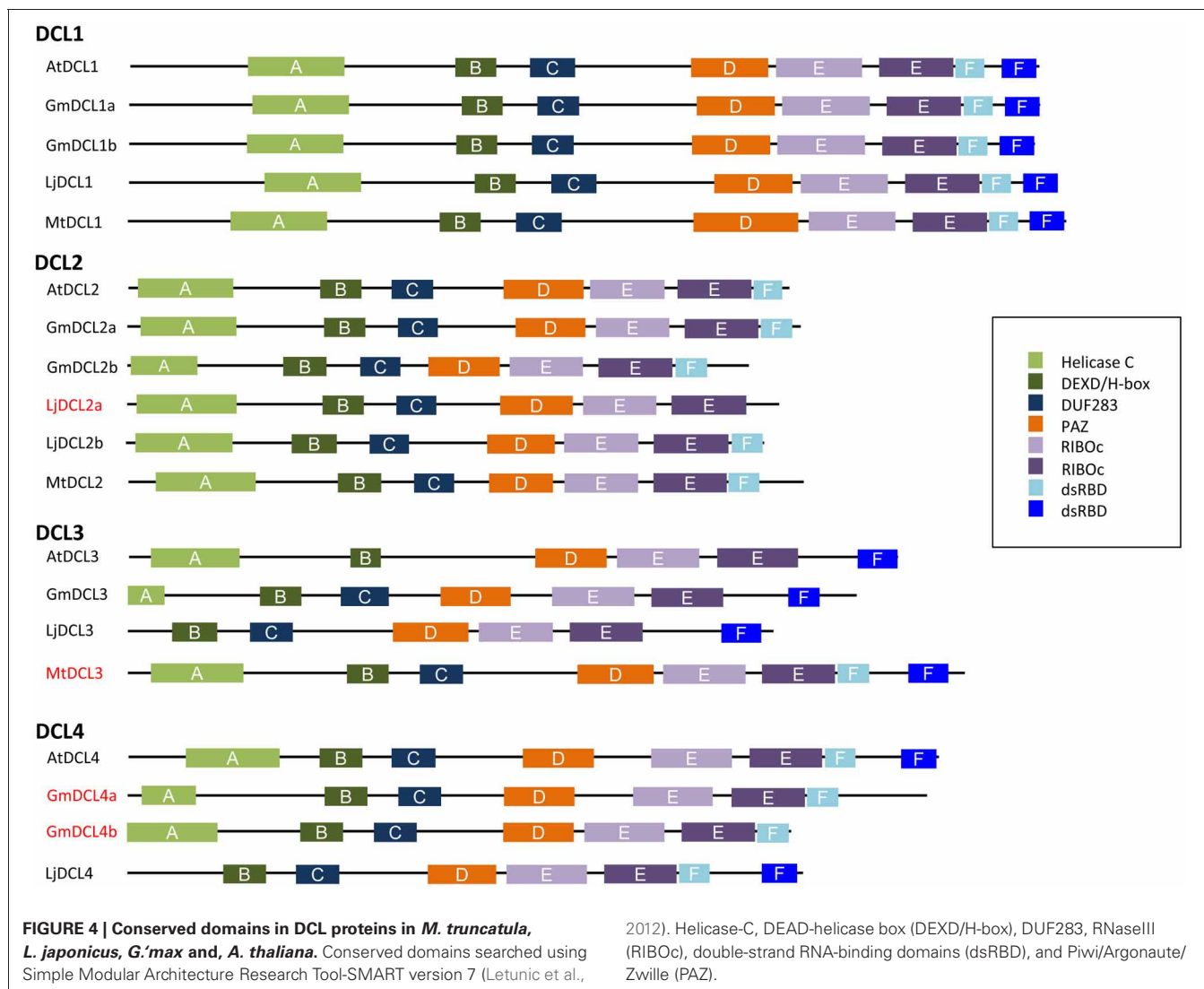
Presence/absence as well as number or type of dsRBD domains have been proposed to be important criteria to determinate the



**FIGURE 3 | Phylogenetic tree of DICER-like proteins in *M. truncatula*, *L. japonicus*, *G. max*, *A. thaliana*, *P. trichocarpa*, and *O. sativa*.** Full length sequences of the predicted DCL proteins from *A. thaliana*, *P. trichocarpa*, and *O. sativa* were retrieved from Genbank (NCBI). DCL from *M. truncatula*, soybean (*G. max*) and *L. japonicus* were searched by tBLASTX in genomic sequence databases: legumes *Lotus japonicus* (Miyakogusa.jp 2.5, Sato et al., 2008), *Medicago truncatula* (Mt 3.5.1, Young et al., 2011), and

*Glycine max* (Glyma1.181, <http://www.plantgdb.org/GmGDB/>), and named according to their similarity with *A. thaliana* proteins or according to previous publications (correspondences given in Data Sheet 3). Phylogenetic trees were constructed thanks to T-Coffee software. The phylogenetic tree was generated with MEGA4.0 software using the Neighbor joining tree (Capitão et al., 2011). The four DCL clades are indicated in the right part of the figure.





substrate specificity and the interaction of AGOs with associated proteins, including DCLs, of the smRNA pathways (Hiraguri et al., 2005; Eamens et al., 2009). Fusion GFP proteins of DCL1, DCL4, and other proteins involved in smRNA pathways like HYPONASTIC LEAVES (HYL1) and DRB (Double-strand RNA Binding) proteins suggested co-localization in specialized ribonucleoproteins in *A. thaliana* (Hiraguri et al., 2005). It has been proposed that dsRBD and PAZ play roles in recognizing and processing the RNA substrates while dsRBD play additional functions by interacting with other proteins of smRNA biogenesis pathways (Hiraguri et al., 2005; Eamens et al., 2009). It will be interesting to investigate the precise function of the *DCL* genes that have been duplicated in some legumes or present different dsRBD compositions in relation to smRNA biogenesis.

#### ARGONAUTES PROTEINS AND THEIR CONSERVED DOMAINS IN LEGUMES: ANOTHER COMPLEX STORY

AGOs are RNA binding proteins that are key effectors of the RISC complexes. AGO-like proteins have been found in bacteria, archaea

and eukaryotes (Hutvagner and Simard, 2008). *A. thaliana*, rice and maize possess 10, 19, and 18 AGO proteins respectively (Kapoor et al., 2008; Vaucheret, 2008; Qian et al., 2011). Similarly to DCLs, certain AGOs show functional redundancy (Vazquez et al., 2010). Plant AGO proteins fall into four distinct clades with different properties in term of smRNA recognition based on sequence similarity: AGO1/10, AGO5, AGO2/3/7, and AGO4/6/8/9 (Mallory and Vaucheret, 2010; Czech and Hannon, 2011). AGO proteins show preferences for the nucleotide placed at the 5' end of the smRNAs as shown by immunoprecipitation of AGO complexes in *Arabidopsis* (Mi et al., 2008; Montgomery et al., 2008; Takeda et al., 2008; Havecker et al., 2010). For instance, AGO1 and AGO10 bind 21 nt (and eventually 22 nt) miRNAs with a 5' uridine; AGO4/6/9 and AGO2/3 prefer 5' adenine residues on 24 and 21 nt siRNA respectively, while AGO5 binds smRNAs with a 5' cytosine (Manavella et al., 2011). However, exceptions to this nucleotide preference may exist and these rules are not absolute. In addition, a specific AGO, AGO7, *Arabidopsis* essentially binds the conserved miR390, which



contains a 5' adenine in *Arabidopsis*. Changing this residue into a cytosine did not produce a difference in AGO7 preference for miR390 (Montgomery et al., 2008) showing that the 5' end is not the only determinant of AGO binding, at least in this case.

Our search by TBLASTX pointed out the presence of 21, 12, and 9 putative AGOs in *G. max*, *M. truncatula* and *L. japonicus*, respectively, as well as 14 homologs in *P. trichocarpa* (Tuskan et al., 2006; *Populus trichocarpa* v3.0, DOE-JGI, <http://www.phytozome.net/poplar>). A previous search in *M. truncatula* (Capitão et al., 2011) already described 12 AGOs based on the MtV3.0 genome (correspondence between ours and theirs AGOs is given in Data Sheet 3). As shown in the AGO phylogenetic tree (Figure 5), legumes AGOs clearly fall into the four clades described in *Arabidopsis*. In soybean, homologs for each AtAGO were found, except AGO8. In fact, the eventual absence of AGO8 is shared by the three legumes, despite that the genomes are not complete. This may be consistent with Takeda et al. (2008) who suggested that AGO8 may be a pseudogene in *A. thaliana*. Several AtAGO homologs could not be identified in present *M. truncatula* and *L. japonicus* genomic databases. For instance, AGO3 and AGO9 are absent both in *Lotus* and *M. truncatula* genomic databases (Capitão et al., 2011; Young et al., 2011). However, according to its strong similarity with the soybean AGO9, we propose to rename MtAGO11a (Capitão et al., 2011) in MtAGO9. In addition, as AGO2 and AGO3 belong to the same clade, it is possible that the absence of AGO3 in the two model legumes is compensated by the presence of an additional AGO2 homolog (two AGO2 homologs instead of AGO2 and AGO3). In *Arabidopsis*, several evidences from *Atago2* mutants pointed out that AGO2 but not AGO3 plays fundamental roles in defense against particular virus infections (Harvey et al., 2011; Wang et al., 2011), suggesting that they may have specialized functions. On the other hand, expression analysis suggest that AGO2 and AGO3 are more expressed in seeds than other tissues and accumulation of smRNA species is similar in *ago2* and *ago3* mutants (Takeda et al., 2008). These latter experiments support that there is functional redundancy of AGO2 and AGO3 in *Arabidopsis*, notably in seeds.

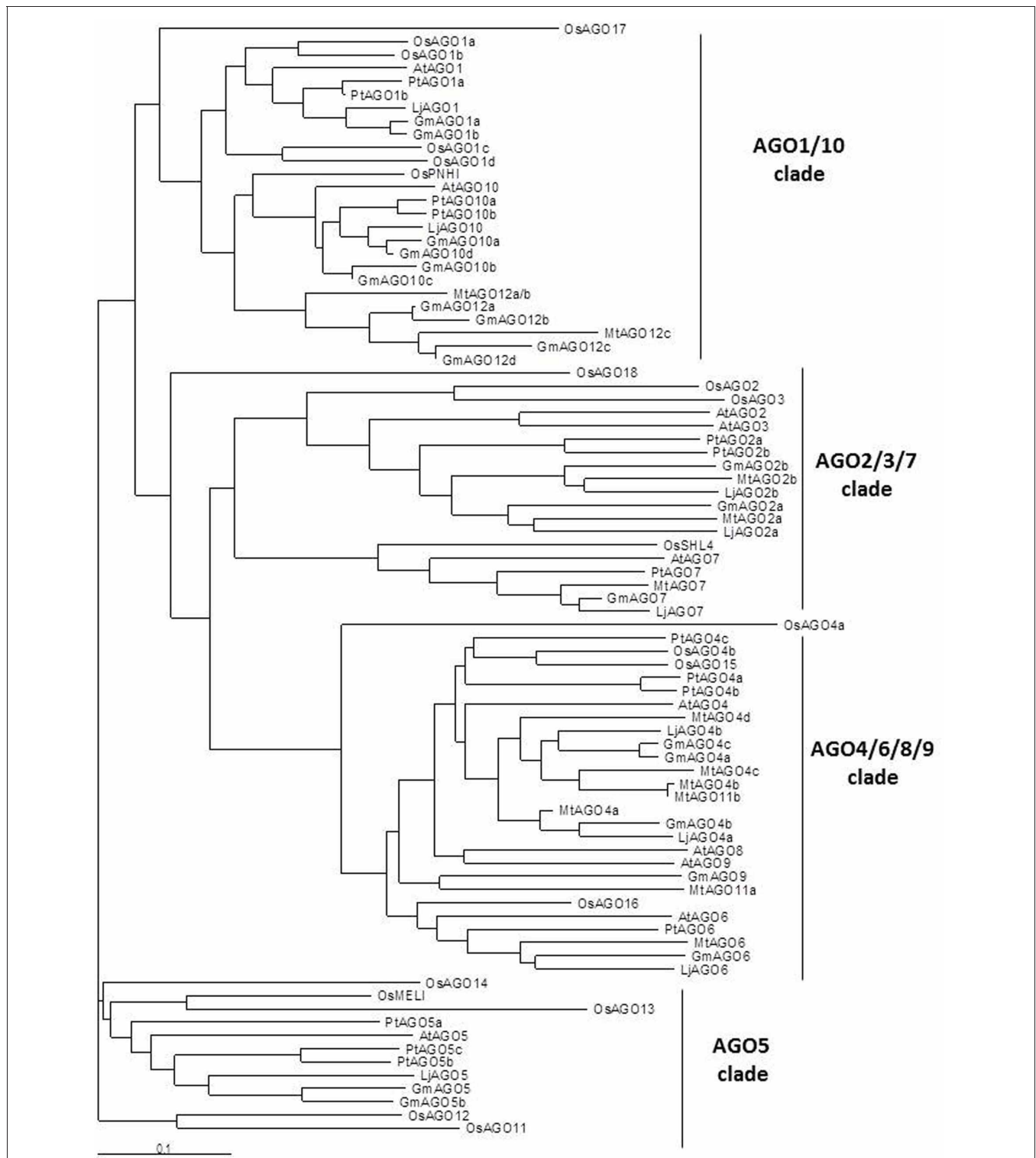
Because of their fundamental roles in miRNA function, absence of clear homologs of AGO1 and AGO10 genes in *M. truncatula* databases was much more striking. In *Arabidopsis*, functional analyses suggested that AGO1 mainly acts through slicing and controls miRNA function. On the other hand, AGO10 acts through translational repression of the miRNA targets and belongs to the same clade (for references see Manavella et al., 2011). As *Atago10* mutants are affected neither in gene silencing nor in smRNA accumulation (Takeda et al., 2008), and *ago1* single mutant and *ago1 ago10* double mutants are embryo lethal (Lynn et al., 1999) it seems that they are not functionally redundant. However, partial functional redundancy between AGO1 and AGO10 proteins has been described in *Arabidopsis* (Maunoury and Vaucheret, 2011). AGO10 apparently preferentially associates with members of the miR165/166 family and some miRNA/miRNA\* structural features seemed to be essential for its activity. AGO10 protein might thus act avoiding miR165/166 loading into AGO1, and hence attenuating the action of these miRNAs by cleavage (Ji et al., 2011; Manavella et al., 2011). In *M. truncatula*, Capitão et al. (2011) first identified an

expressed sequence contig (TC126810, TIGR MtGi9.0) encoding a putative AGO1 homolog. However, in the last version of TIGR (MtGi11.0, March 2011), this TC was split into two contigs (TC188472 and TC194233), which in fact correspond to the so-called MtAGO12a/b genes. Although these AGO12 isoforms fall into the AGO1/AGO10 clade, they are more closely related to AGO10 (Figure 5). Thus, the absence of AGO1, but also of AGO5, in *M. truncatula* databases remains intriguing and may be due to the lack of completion of the *M. truncatula* genome.

Vaucheret (2008) and Kapoor et al. (2008) reported that new AGO genes arise from gene duplication events. This type of gene diversification clearly appears in the phylogenetic tree which shows that, for many AtAGOs, more than one homolog was found in nearly all plant species we compared. Looking at the phylogenetic relationships inside the AGO4/6/8/9 clade, AGO4 diversification was important in all species with 2–4 homologs. In addition, AGO4s from the three model legumes appeared clearly separated from the corresponding homologs in non-legumes. A divergence in smRNA biogenesis proteins has been observed in vertebrates, invertebrates and plants suggesting a lineage-specific modification of gene regulation by smRNAs, and pointing out to the plasticity of genome for the evolution of novel regulatory networks (Murphy et al., 2008). In this case, the possibility of an AGO4 evolution in legumes toward an adaptation of regulatory smRNA network and symbiotic interactions may be worth considering.

To further analyse putative functions of AGOs, we compared their conserved domains: MID, PIWI, PAZ, and DUF1787 domain of unknown function (Vaucheret, 2008). These domains are linked to different activities of AGOs: the MID domain binds the smRNA 5' end (Wang et al., 2009; Parker, 2010); The PIWI domain is responsible of the catalytic activity of AGO proteins (Baumberger and Baulcombe, 2005); the PAZ domain is necessary to recognize the 3' end of the smRNA for loading into the RISC and determines the AGO specificity. Among the identified domains in legume AGOs (Data Sheet 4), only some appeared unusual. For instance, GmAGO10c and MtAGO4a lacked PAZ and DUF1785 domains, suggesting that they are pseudogenes. On the other hand, MtAGO12c (Medtr2g059590, Capitão et al., 2011) had no PIWI domain. For that, we assume that this protein, even if it also belongs to the AGO1/AGO10 clade, may certainly not play similar functions than AGO1. Again, more precise analyses of the corresponding genomic regions and completion of the genome sequencing are required to support the absence of these sequences as errors in gene annotation due to the presence of very long introns for instance may explain some of the difficulties to detect gene homologs in legumes.

Finally, we decided to compare more precisely the sequence of the PIWI domains in legume AGOs (Data Sheet 5). Indeed, activity of the PIWI domain has been associated to the presence of conserved catalytic residues, in particular three metal-chelating amino acid residues Asp-Asp-His (DDH). Lack of one of these residues has been reported to alter AGO slicing but not its binding activity (Qi et al., 2006; Kapoor et al., 2008; Nowotny and Yang, 2009). For instance, AGO4 proteins with a modified DDH motif still bind siRNAs but lose their endonuclease activity (Qi et al., 2006). However, in *Arabidopsis*, the His residue of AGO2



**FIGURE 5 | Phylogenetic tree of AGO proteins in *M. truncatula*, *L. japonicus*, *G. max*, *A. thaliana*, *P. trichocarpa*, and *O. sativa*.** Full length sequences of the predicted AGO proteins from *A. thaliana*, *P. trichocarpa* and *O. sativa* were retrieved from Genbank (NCBI). DCL from *M. truncatula*, soybean (*G. max*) and *L. japonicus* were searched by tBLASTX in genomic sequence databases: legumes *Lotus japonicus* (Miyakogusa.jp 2.5, Sato et al., 2008), *Medicago truncatula* (Mt 3.5.1, Young et al., 2011),

and *Glycine max* (Glyma1.181, <http://www.plantgdb.org/GmGDB/>), and named according to their similarity with *A. thaliana* proteins or according to previous publications (correspondances given in Data Sheet 3). Phylogenetic trees were constructed thanks to T-Coffee software. The phylogenetic tree was generated with MEGA4.0 software using the Neighbor joining tree (Capitão et al., 2011). The four AGO clades are indicated in the right part of the figure.

and AGO3 may be replaced by Asp (DDD) without affecting their endonuclease activity (Baumberger and Baulcombe, 2005; Montgomery et al., 2008) and the same happened for 9 AGO proteins in rice (Kapoor et al., 2008). As shown in Data Sheet 1, all legume AGOs with a PIWI domain present either a DDH or a DDD motif, like in *Arabidopsis*. In addition, Baumberger and Baulcombe (2005) reported that a conserved Histidine residue at position +800 in the PIWI domain of AGO1 was required for its endonuclease activity, at least *in-vitro*. As expected, according to their role in RNA-dependent DNA methylation rather than RNA cleavage, AtAGO4/6/8/9 proteins contain another amino-acid residue, an alanine (A), a serine (S) or a proline (P), at this conserved position. We thus looked for the presence of the conserved H<sub>800</sub> residue in legume AGOs. Like in *A. thaliana*, all AGO4, AGO6 and AGO9 proteins had no H<sub>800</sub>. Unexpectedly, the same is true in MtAGO12a/b proteins, again suggesting that they may not replace AGO1 cleavage function in *M. truncatula*, but may rather play similar translational regulatory functions to AGO10 (Zhu et al., 2011). Interestingly, AGO2b proteins of the three model legumes present an asparagine residue (N) instead of the H. As this amino-acid change did not occur in *Arabidopsis*, we wondered whether this substitution was found in other non-legumes species. Blasting the PIWI domain of AtAGO2 with the corresponding domains of AGO2 in rice and poplar, no changes in H<sub>800</sub> was observed. In addition, when we analyzed all AGOs in both species, the H<sub>800</sub> residue was never replaced by a N, although other amino-acids may be found such as Cysteine (C) in OsAGO17 or Glutamine (Q) in PtAGO4c. The presence of this N residue at a conserved key position of AGO2b proteins thus suggests a specific but conserved role of this isoform in legumes.

## CONCLUDING REMARKS

Since five years, the rapid progress of genomic technologies has allowed the construction and high-throughput sequencing of a very large set of smRNA libraries in model legumes, in

particular from *M. truncatula* and *G. max*. Compilation of data from miRBase and smRNA deep sequencings revealed the largest set of novel miRNA families ever observed in a plant family. Few conserved and legume-specific miRNAs have been functionally studied. Further analysis of the legume miRNAs as well as phased siRNAs and their putative targets will be essential to better understand the role of smRNAs in the control of key processes during nodule development, nitrogen fixation and defense against pathogens.

On the other hand, DCL and AGO proteins are key players in smRNA biogenesis and functions, which are highly conserved processes in plants. In legumes, many DCL and AGO proteins, like DCL2, AGO2, AGO4, AGO10 present a huge diversification, in comparison to *Arabidopsis thaliana*. However, in some legume models, highly conserved plant DCL and AGO formshave not been identified until now (e.g., MtDCL4 or MtAGO1/AGO10) or present a non-conventional domain composition. Further exhaustive genome analysis may be necessary to confirm the absence of those genes. Further analysis of DCL and AGO isoforms in legumes are necessary to clarify whether duplication of some isoforms could be correlated with a specialization in function and/or expression in legumes, in particular during symbiotic or pathogenic interactions.

## ACKNOWLEDGMENTS

Pilar Bustos-Sanmamed was the recipient of a Marie Curie Postdoctoral IE Fellowship (European Commission, MEDEPIMIR, N° PIEF-GA-2010-273743). Jérémie Bazinwork was supported by a post-doctoral contract from University Paris Diderot Paris 7, France. This work was supported by the program Saclay Plant Sciences (SPS, ANR-10-LABX-40).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: [http://www.frontiersin.org/Plant\\_Genetics\\_and\\_Genomics/10.3389/fpls.2013.00236/abstract](http://www.frontiersin.org/Plant_Genetics_and_Genomics/10.3389/fpls.2013.00236/abstract)

## REFERENCES

- Allen, E., and Howell, M. C. (2010). miRNAs in the biogenesis of trans-acting siRNAs in higher plants. *Semin. Cell Dev. Biol.* 21, 798–804. doi: 10.1016/j.semcdb.2010.03.008
- Allen, E., Xie, Z., Gustafson, A. M., Sung, G. H., Spatafora, J. W., and Carrington, J. C. (2004). Evolution of microRNA genes by inverted duplication of target gene sequences in *Arabidopsis thaliana*. *Nat. Genet.* 36, 1282–1290. doi: 10.1038/ng1478
- Allen, E., Xie, Z., Gustafson, A. M., and Carrington, J. C. (2005). microRNA-directed phasing during trans-acting siRNA biogenesis in plants. *Cell* 121, 207–221.
- Ariel, F., Brault-Hernandez, M., Laffont, C., Huault, E., Brault, M., Plet, J., et al. (2012). Two direct targets of cytokinin signaling regulate symbiotic nodulation in *Medicago truncatula*. *Plant Cell* 24, 3838–3852. doi: 10.1105/tpc.112.103267
- Baumberger, N., and Baulcombe, D. C. (2005). *Arabidopsis* ARGONAUTE1 is an RNA slicer that selectively recruits microRNAs and short interfering RNAs. *Proc. Natl. Acad. Sci. U.S.A.* 102, 11928–11933. doi: 10.1073/pnas.0505461102
- Bazin, J., Bustos-Sanmamed, P., Hartmann, C., Lelandais-Brière, C., and Crespi, M. (2012). Complexity of miRNA-dependent regulation in root symbiosis. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 1570–1579. doi: 10.1098/rstb.2011.0228
- Ben Amor, B., Wirth, S., Merchan, F., Laporte, P., d'Aubenton-Carafa, Y., Hirsch, J., et al. (2009). Novel long non-protein coding RNAs involved in *Arabidopsis* differentiation and stress responses. *Genome Res.* 19, 57–69. doi: 10.1101/gr.080275.108
- Boualem, A., Laporte, P., Jovanovic, M., Laffont, C., Plet, J., Combier, J. P., et al. (2008). MicroRNA166 controls root and nodule development in *Medicago truncatula*. *Plant J.* 54, 876–887. doi: 10.1111/j.1365-313X.2008.03448.x
- Bourcy, M., Berrabah, F., Ratet, P., and Gourion, B. (2013). Failure of self-control: defense like reactions during legume/rhizobia symbiosis. *Plant Signal. Behav.* 8, pii:e23915.
- Bove, J., Hord, C. L., and Mullen, M. A. (2006). The blossoming of RNA biology: novel insights from plant systems. *RNA* 12, 2035–2046. doi: 10.1261/rna.303806
- Branscheid, A., Devers, E. A., May, P., and Krajinski, F. (2011). Distribution pattern of small RNA and degradome reads provides information on miRNA gene structure and regulation. *Plant Signal. Behav.* 6, 1609–1611. doi: 10.4161/psb.6.10.17305
- Brodersen, P., Sakvarelidze-Achard, L., Bruun-Rasmussen, M., Dunoyer, P., Yamamoto, Y. Y., Sieburth, L., et al. (2008). Widespread translational inhibition by plant miRNAs and siRNAs. *Science* 320, 1185–1190. doi: 10.1126/science.1159151
- Capitão, C., Paiva, J. A., Santos, D. M., and Fevereiro, P. (2011). In *Medicago truncatula*, water deficit modulates the transcript accumulation of components of small RNA pathways. *BMC Plant Biol.* 11:79. doi: 10.1186/1471-2229-11-79
- Carmell, M. A., and Hannon, G. J. (2004). RNase III enzymes and the initiation of gene silencing. *Nat.*

- Struct. Mol. Biol.* 11, 214–218. doi: 10.1038/nsmb729
- Cerutti, H., and Casas-Mollano, J. A. (2006). On the origin and functions of RNA-mediated silencing: from protists to man. *Curr. Genet.* 50, 81–99. doi: 10.1007/s00294-006-0078-x
- Chapman, E. J., and Carrington, J. C. (2007). Specialization and evolution of endogenous small RNA pathways. *Nat. Rev. Genet.* 8, 884–896. doi: 10.1038/nrg2179
- Chen, H. M., Chen, L. T., Patel, K., Li, Y. H., Baulcombe, D. C., and Wu, S. H. (2010). 22-nucleotide RNAs trigger secondary siRNA biogenesis in plants. *Proc. Natl. Acad. Sci. U.S.A.* 107, 15269–15274. doi: 10.1073/pnas.1001738107
- Combiér, J. P., Frugier, F., de Billy, F., Boualem, A., El-Yahyaoui, F., Moreau, S., et al. (2006). MtHAP2-1 is a key transcriptional regulator of symbiotic nodule development regulated by microRNA169 in *Medicago truncatula*. *Genes Dev.* 20, 3084–3088. doi: 10.1101/gad.402806
- Corradi, N., and Bonfante, P. (2012). The arbuscular mycorrhizal symbiosis: origin and evolution of a beneficial plant infection. *PLoS Pathog.* 8:e1002600. doi: 10.1371/journal.ppat.1002600
- Cuperus, J. T., Fahlgren, N., and Carrington, J. C. (2011). Evolution and functional diversification of MIRNA genes. *Plant Cell* 23, 431–442. doi: 10.1105/tpc.110.082784
- Curtin, S. J., Kantar, M. B., Yoon, H. W., Whaley, A. M., Schlueter, J. A., and Stupar, R. M. (2012). Co-expression of soybean dicer-like genes in response to stress and development. *Funct. Integr. Genomics* 12, 671–682. doi: 10.1007/s10142-012-0278-z
- Czech, B., and Hannon, G. J. (2011). Small RNA sorting: matchmaking for argonautes. *Nat. Rev. Genet.* 12, 19–31. doi: 10.1038/nrg2916
- Deinum, E. E., Geurts, R., Bisseling, T., and Mulder, B. M. (2012). Modeling a cortical auxin maximum for nodulation: different signatures of potential strategies. *Front. Plant Sci.* 3:96. doi: 10.3389/fpls.2012.00096
- De Luis, A., Markmann, K., Cognat, V., Holt, D. B., Charpentier, M., Parniske, M., et al. (2012). Two microRNAs linked to nodule infection and nitrogen-fixing ability in the legume *Lotus japonicus*. *Plant Physiol.* 160, 2137–2154. doi: 10.1104/pp.112.204883
- Devers, E. A., Branscheid, A., May, P., and Krajinski, F. (2011). Stars and symbiosis: microRNA- and microRNA\*-mediated transcript cleavage involved in arbuscular mycorrhizal symbiosis. *Plant Physiol.* 156, 1990–2010. doi: 10.1104/pp.111.172627
- D'Haeseleer, K., Den Herder, G., Laffont, C., Plet, J., Mortier, V., Lelandais-Brière, C., et al. (2011). Transcriptional and post-transcriptional regulation of a NAC1 transcription factor in *Medicago truncatula* roots. *New Phytol.* 191, 647–661. doi: 10.1111/j.1469-8137.2011.03719.x
- Dunoyer, P., Himber, C., Ruiz-Ferrer, V., Alioua, A., and Voinnet, O. (2007). Intra- and intercellular RNA interference in *Arabidopsis thaliana* requires components of the microRNA and heterochromatic silencing pathways. *Nat. Genet.* 39, 848–856. doi: 10.1038/ng2081
- Eamens, A. L., Smith, N. A., Curtin, S. J., Wang, M. B., and Waterhouse, P. M. (2009). The *Arabidopsis thaliana* double-stranded RNA binding protein DRB1 directs guide strand selection from microRNA duplexes. *RNA* 15, 2219–2235. doi: 10.1261/rna.1646909
- Gascioli, V., Mallory, A. C., Bartel, D. P., and Vaucheret, H. (2005). Partially redundant functions of *Arabidopsis* dicer-like enzymes and a role for DCL4 in producing trans-acting siRNAs. *Curr. Biol.* 15, 1494–1500. doi: 10.1016/j.cub.2005.07.024
- German, M. A., Luo, S., Schroth, G., Meyers, B. C., and Green, P. J. (2009). Construction of parallel analysis of RNA ends (PARE) libraries for the study of cleaved miRNA targets and the RNA degradome. *Nat. Protoc.* 4, 356–362. doi: 10.1038/nprot.2009.8
- Gill, N., Findley, S., Walling, J. G., Hans, C., Ma, J., Doyle, J., et al. (2009). Molecular and chromosomal evidence for allopolyploidy in soybean. *Plant Physiol.* 151, 1167–1174. doi: 10.1104/pp.109.137935
- Griffiths-Jones, S., Grocock, R. J., van Dongen, S., Bateman, A., and Enright, A. J. (2006). miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res.* 34, D140–D144. doi: 10.1093/nar/gkj112
- Gu, M., Xu, K., Chen, A., Zhu, Y., Tang, G., and Xu, G. (2010). Expression analysis suggests potential roles of microRNAs for phosphate and arbuscular mycorrhizal signaling in *Solanum lycopersicum*. *Physiol. Plant.* 138, 226–237. doi: 10.1111/j.1399-3054.2009.01320.x
- Harvey, J. J., Lewsey, M. G., Patel, K., Westwood, J., Heimstädt, S., Carr, J. P., et al. (2011). An antiviral defense role of AGO2 in plants. *PLoS ONE* 6:e14639. doi: 10.1371/journal.pone.0014639
- Havecker, E. R., Wallbridge, L. M., Hardcastle, T. J., Bush, M. S., Kelly, K. A., Dunn, R. M., et al. (2010). The *Arabidopsis* RNA-directed DNA methylation argonautes functionally diverge based on their expression and interaction with target loci. *Plant Cell* 22, 321–334. doi: 10.1105/tpc.109.072199
- Henderson, I. R., Zhang, X., Lu, C., Johnson, L., Meyers, B. C., Green, P. J., et al. (2006). Dissecting *Arabidopsis thaliana* dicerfunction in small RNA processing, gene silencing and DNA methylation patterning. *Nat. Genet.* 38, 721–725. doi: 10.1038/ng1804
- Hiraguri, A., Itoh, R., Kondo, N., Nomura, Y., Aizawa, D., Murai, Y., et al. (2005). Specific interactions between dicer-like proteins and HYL1/DRB-family dsRNA-binding proteins in *Arabidopsis thaliana*. *Plant Mol. Biol.* 57, 173–188. doi: 10.1007/s11103-004-6853-5
- Hutvagner, G., and Simard, M. J. (2008). Argonaute proteins: key players in RNA silencing. *Nat. Rev. Mol. Cell Biol.* 9, 22–32. doi: 10.1038/nrm2321
- Jagadeeswaran, G., Zheng, Y., Li, Y. F., Shukla, L. I., Matts, J., Hoyt, P., et al. (2009). Cloning and characterization of small RNAs from *Medicago truncatula* reveals four novel legume-specific microRNA families. *New Phytol.* 184, 85–98. doi: 10.1111/j.1469-8137.2009.02915.x
- Ji, L., Liu, X., Yan, J., Wang, W., Yumul, R. E., Kim, Y. J., et al. (2011). ARGONAUTE10 and ARGONAUTE1 regulate the termination of floral stem cells through twomicroRNAs in *Arabidopsis*. *PLoS Genet.* 7:e1001358. doi: 10.1371/journal.pgen.1001358
- Jones, J. D., and Dangl, J. L. (2006). The plant immune system. *Nature* 444, 323–329. doi: 10.1038/nature05286
- Jouannet, V., Moreno, A. B., Elmayan, T., Vaucheret, H., Crespi, M. D., and Maizel, A. (2012). Cytoplasmic *Arabidopsis* AGO7 accumulates in membrane-associated siRNA bodies and is required for ta-siRNA biogenesis. *EMBO J.* 31, 1704–1713. doi: 10.1038/emboj.2012.20
- Kapoor, M., Arora, R., Lama, T., Nijhawan, A., Khurana, J. P., Tyagi, A. K., et al. (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. doi: 10.1186/1471-2164-9-451
- Kawashima, C. G., Yoshimoto, N., Maruyama-Nakashita, A., Tsuchiya, Y. N., Saito, K., Takahashi, H., et al. (2009). Sulphur starvation induces the expression of microRNA-395 and one of its target genes but in different cell types. *Plant J.* 57, 313–321. doi: 10.1111/j.1365-313X.2008.03690.x
- Khan, G. A., Declerck, M., Sorin, C., Hartmann, C., Crespi, M., and Lelandais-Brière, C. (2011). MicroRNAs as regulators of root development and architecture. *Plant Mol. Biol.* 77, 47–58. doi: 10.1007/s11103-011-9793-x
- Khraiwesh, B., Zhu, J. K., and Zhu, J. (2012). Role of miRNAs and siRNAs in biotic and abiotic stress responses of plants. *Biochim. Biophys. Acta* 1819, 137–148. doi: 10.1016/j.bbagr.2011.05.001
- Kulcheski, F. R., de Oliveira, L. F., Molina, L. G., Almerão, M. P., Rodrigues, F. A., Marcolino, J., et al. (2011). Identification of novel soybean microRNAs involved in abiotic and biotic stresses. *BMC Genomics* 12:307. doi: 10.1186/1471-2164-12-307
- Lauressergues, D., Delaux, P. M., Formey, D., Lelandais-Brière, C., Fort, S., Cottaz, S., et al. (2012). The microRNA miR171h modulates arbuscular mycorrhizal colonization of *Medicago truncatula* by targeting NSP2. *Plant J.* 72, 512–522. doi: 10.1111/j.1365-313X.2012.05099.x
- Lelandais-Brière, C., Naya, L., Sallet, E., Calenge, F., Frugier, F., Hartmann, C., et al. (2009). Genome-wide *Medicago truncatula* small RNA analysis revealed novel microRNAs and isoforms differentially regulated in roots and nodules. *Plant Cell* 21, 2780–2796. doi: 10.1105/tpc.109.068130
- Letunic, I., Doerks, T., and Bork, P. (2012). SMART 7: recent updates to the protein domain annotation resource. *Nucleic Acids Res.* 40, D302–D305. doi: 10.1093/nar/gkr931
- Li, F., Pignatta, D., Bendix, C., Brunkard, J. O., Cohn, M. M., Tung, J., et al. (2012). MicroRNA regulation of plant innate immune receptors. *Proc. Natl. Acad. Sci. U.S.A.* 109, 1790–1795. doi: 10.1073/pnas.1118282109



- Li, H., Deng, Y., Wu, T., Subramanian, S., and Yu, O. (2010). Misexpression of miR482, miR1512, and miR1515 increases soybean nodulation. *Plant Physiol.* 153, 1759–1770. doi: 10.1104/pp.110.156950
- Li, H., Dong, Y., Yin, H., Wang, N., Yang, J., Liu, X., et al. (2011). Characterization of the stress associated microRNAs in glycine max by deep sequencing. *BMC Plant Biol.* 11:170. doi: 10.1186/1471-2229-11-170
- Liang, G., He, H., and Yu, D. (2012). Identification of nitrogen starvation-responsive microRNAs in *Arabidopsis thaliana*. *PLoS ONE* 7:e48951. doi:10.1371/journal.pone.0048951
- Liang, G., Yang, F., and Yu, D. (2010). MicroRNA395 mediates regulation of sulfate accumulation and allocation in *Arabidopsis thaliana*. *Plant J.* 62, 1046–1057.
- Liu, B., Li, P., Li, X., Liu, C., Cao, S., Chu, C., et al. (2005). Loss of function of OsDCL1 affects microRNA accumulation and causes developmental defects in rice. *Plant Physiol.* 139, 296–305. doi:10.1104/pp.105.063420
- Liu, Q., Feng, Y., and Zhu, Z. (2009). Dicer-like (DCL) proteins in plants. *Funct. Integr. Genomics* 9, 277–286. doi:10.1007/s10142-009-0111-5
- López-Gómez, M., Sandal, N., Stougaard, J., and Boller, T. (2012). Interplay of flg22-induced defence responses and nodulation in *Lotus japonicus*. *J. Exp. Bot.* 63, 393–401. doi: 10.1093/jxb/err291
- Lu, C., Kulkarni, K., Souret, F. F., MuthuValliappan, R., Tej, S. S., Poethig, R. S., et al. (2006). MicroRNAs and other small RNAs enriched in the *Arabidopsis* RNA-dependent RNA polymerase-2 mutant. *Genome Res.* 16, 1276–1288. doi:10.1101/gr.5530106
- Lynn, K., Fernandez, A., Aida, M., Sedbrook, J., Tasaka, M., Masson, P., et al. (1999). The PINHEAD/ZWILLE gene acts pleiotropically in *Arabidopsis* development and has overlapping functions with the ARGONAUTE1 gene. *Development* 126, 469–481.
- Mallory, A., and Vaucheret, H. (2010). Form, function, and regulation of ARGONAUTE proteins. *Plant Cell* 22, 3879–3889. doi: 10.1105/tpc.110.080671
- Manavella, P. A., Koenig, D., and Weigel, D. (2012). Plant secondary siRNA production determined by microRNA-duplex structure. *Proc. Natl. Acad. Sci. U.S.A.* 109, 2461–2466. doi: 10.1073/pnas.1200169109
- Manavella, P. A., Weigel, D., and Wu, L. (2011). Argonaute10 as a miRNA locker. *Cell* 145, 173–174. doi: 10.1016/j.cell.2011.03.045
- Maunoury, N., and Vaucheret, H. (2011). AGO1 and AGO2 act redundantly in miR408-mediated Plantacyanin regulation. *PLoS ONE* 6:e28729. doi: 10.1371/journal.pone.0028729
- Margis, R., Fusaro, A. F., Smith, N. A., Curtin, S. J., Watson, J. M., Finnegan, E. J., et al. (2006). The evolution and diversification of dicers in plants. *FEBS Lett.* 580, 2442–2450. doi: 10.1016/j.febslet.2006.03.072
- Marin, E., Jouannet, V., Herz, A., Lokere, A. S., Weijers, D., Vaucheret, H., et al. (2010). miR390, *Arabidopsis* TAS3 tasiRNAs, and their auxin response factor targets define an autoregulatory network quantitatively regulating lateral root growth. *Plant Cell* 22, 1104–1117. doi: 10.1105/tpc.109.072553
- Marchetti, M., Capela, D., Glew, M., Cruveilles, S., Chane-Woon-Ming, B., Girs, C., et al. (2010). Experimental evolution of a plant pathogen into a legume symbiont. *PLoS Biol.* 8:e1000280. doi: 10.1371/journal.pbio.1000280
- Matamoros, M. A., Baird, L. M., Escuredo, P. R., Dalton, D. A., Minchin, F. R., Iturbe-Ormaetxe, I., et al. (1999). Stress-induced legume root nodule senescence. Physiological, biochemical, and structural alterations. *Plant Physiol.* 121, 97–112. doi: 10.1104/pp.121.1.97
- Mendoza-Soto, A. B., Sánchez, F., and Hernández, G. (2012). MicroRNAs as regulators in plant metal toxicity response. *Front. Plant Sci.* 3:105. doi: 10.3389/fpls.2012.00105
- Meyers, B. C., Axtell, M. J., Bartel, B., Bartel, D. P., Baulcombe, D., Bowman, J. L., et al. (2008). Criteria for annotation of plant MicroRNAs. *Plant Cell* 20, 3186–3190. doi: 10.1105/tpc.108.064311
- Mi, S., Cai, T., Hu, Y., Chen, Y., Hodges, E., Ni, F., et al. (2008). Sorting of small RNAs into *Arabidopsis* argonaute complexes is directed by the 5' terminal nucleotide. *Cell* 133, 116–127. doi: 10.1016/j.cell.2008.02.034
- Montgomery, T. A., Howell, M. D., Cuperus, J. T., Li, D., Hansen, J. E., Alexander, A. L., et al. (2008). Specificity of ARGONAUTE7-miR390 interaction and dual functionality in TAS3 trans-acting siRNA formation. *Cell* 133, 128–141. doi: 10.1016/j.cell.2008.02.033
- Murphy, D., Dancis, B., Brown, J. R. (2008). The evolution of core proteins involved in microRNA biogenesis. *BMC Evol. Biol.* 8:92. doi: 10.1186/1471-2148-8-92
- Naya, L., Khan, G. A., Sorin, C., Hartmann, C., Crespi, M., and Lelandais-Brière, C. (2010). Cleavage of a non-conserved target by a specific miR156 isoform in root apices of *Medicago truncatula*. *Plant Signal. Behav.* 5, 328–331. doi: 10.4161/psb.5.3.11190
- Nowotny, M., and Yang, W. (2009). Structural and functional modules in RNA interference. *Curr. Opin. Struct. Biol.* 19, 286–293. doi: 10.1016/j.sbi.2009.04.006
- Ochman, H., and Moran, M. A. (2001). Gene lost and gene found: evolution of bacterial pathogenesis and symbiosis. *Science* 292, 1096–1099. doi: 10.1126/science.1058543.
- Olmedo-Monfil, V., Durán-Figueroa, N., Arteaga-Vázquez, M., Demesa-Arévalo, E., Autran, D., Grimanelli, D., et al. (2010). Control of female gamete formation by a small RNA pathway in *Arabidopsis*. *Nature* 464, 628–632. doi: 10.1038/nature08828
- Onodera, Y., Haag, J. R., Ream, T., Costa Nunes, P., Pontes, O., and Pikaard, C. S. (2005). Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. *Cell* 120, 613–622. doi: 10.1016/j.cell.2005.02.007
- Parent, J. S., Martínez de Alba, A. E., and Vaucheret, H. (2012). The origin and effect of small RNA signaling in plants. *Front. Plant Sci.* 3:179. doi: 10.3389/fpls.2012.00179
- Parker, J. S. (2010). How to slice: snapshots of Argonaute in action. *Silence* 1:3. doi: 10.1186/1758-907X-1-3
- Peleg-Grossman, S., Melamed-Book, N., and Levine, A. (2013). ROS production during symbiotic infection suppresses pathogenesis-related gene expression. *Plant Signal. Behav.* 7, 409–415. doi: 10.4161/psb.19217
- Peragine, A., Yoshikawa, M., Wu, G., Albrecht, H. L., and Poethig, R. S. (2004). SGS3 and SGS2/SDE1/RDR6 are required for juvenile development and the production of trans-acting siRNAs in *Arabidopsis*. *Genes Dev.* 18, 2368–2379. doi: 10.1101/gad.1231804
- Qi, Y., He, X., Wang, X. J., Kohany, O., Jurka, J., and Hannon, G. J. (2006). Distinct catalytic and non-catalytic roles of ARGONAUTE4 in RNA-directed DNA methylation. *Nature* 443, 1008–1012. doi: 10.1038/nature05198
- Qian, Y., Cheng, Y., Cheng, X., Jiang, H., Zhu, S., and Cheng, B. (2011). Identification and characterization of dicer-like, Argonaute and RNA-dependent RNA polymerase gene families in maize. *Plant Cell Rep.* 30, 1347–1363. doi: 10.1007/s00299-011-1046-6
- Radwan, O., Liu, Y., and Clough, S. J. (2011). Transcriptional analysis of soybean root response to *Fusarium virguliforme*, the causal agent of sudden death syndrome. *Mol. Plant Microbe Interact.* 24, 958–972. doi: 10.1094/MPMI-11-10-0271
- Rajagopalan, R., Vaucheret, H., Trejo, J., and Bartel, D. P. (2006). A diverse and evolutionarily fluid set of microRNAs in *Arabidopsis thaliana*. *Genes Dev.* 20, 3407–3425. doi: 10.1101/gad.1476406
- Rubio, M. C., Becana, M., Sato, S., James, E. K., Tabata, S., and Spaink, H. P. (2007). Characterization of genomic clones and expression analysis of the three types of superoxide dismutases during nodule development in *Lotus japonicus*. *Mol. Plant Microbe Interact.* 20, 262–275. doi: 10.1094/MPMI-20-3-0262
- Sato, S., Nakamura, Y., Kaneko, T., Asamizu, E., Kato, T., Nakao, M., et al. (2008). Genome structure of the legume, *Lotus japonicus*. *DNA Res.* 15, 227–239. doi: 10.1093/dnares/dsn008
- Schmutz, J., Cannon, S. B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., et al. (2010). Genome sequence of the palaeopolyploid soybean. *Nature* 463, 178–183. doi: 10.1038/nature08670
- Shivaprasad, P. V., Chen, H. M., Patel, K., Bond, D. M., Santos, B. A., and Baulcombe, D. C. (2012). A microRNA superfamily regulates nucleotide binding site-leucine-rich repeats and other mRNAs. *Plant Cell* 24, 859–874. doi: 10.1105/tpc.111.095380
- Simon, S. A., Meyers, B. C., and Sherrier, D. J. (2009). MicroRNAs in the rhizobia legume symbiosis. *Plant Physiol.* 151, 1002–1008. doi: 10.1104/pp.109.144345
- Song, Q. X., Liu, Y. F., Hu, X. Y., Zhang, W. K., Ma, B., Chen, S. Y., et al. (2011). Identification of miRNAs and their target genes in developing soybean seeds by deep sequencing. *BMC Plant Biol.* 11:5. doi: 10.1186/1471-2229-11-5
- Song, X., Li, P., Zhai, J., Zhou, M., Ma, L., Liu, B., et al. (2012). Roles of DCL4 and DCL3b in rice phased small RNA biogenesis. *Plant J.* 69, 462–474. doi: 10.1111/j.1365-313X.2011.04805.x



- Soto, M. J., Domínguez-Ferreras, A., Pérez-Mendoza, D., Sanjuán, J., and Olivares, J. (2009). Mutualism versus pathogenesis: the give-and-take in plant-bacteria interactions. *Cell. Microbiol.* 11, 381–388. doi: 10.1111/j.1462-5822.2009.01282.x
- Subramanian, S., Fu, Y., Sunkar, R., Barbazuk, W. B., Zhu, J. K., and Yu, O. (2008). Novel and nodulation-regulated microRNAs in soybean roots. *BMC Genomics* 9:160. doi: 10.1186/1471-2164-9-160
- Subramanian, S., Stacey, G., and Yu, O. (2007). Distinct, crucial roles of flavonoids during legume nodulation. *Trends Plant Sci.* 12, 282–285. doi: 10.1016/j.tplants.2007.06.006
- Sunkar, R., and Jagadeeswaran, G. (2008). *In silico* identification of conserved microRNAs in large number of diverse plant species. *BMC Plant Biol.* 8:37. doi: 10.1186/1471-2229-8-37
- Szittyá, G., Moxon, S., Santos, D. M., Jing, R., Fevereiro, M. P., Moulton, V. T., et al. (2008). High-throughput sequencing of *Medicago truncatula* short RNAs identifies eight new miRNA families. *BMC Genomics* 9:593. doi: 10.1186/1471-2164-9-593
- Takeda, A., Iwasaki, S., Watanabe, T., Utsumi, M., and Watanabe, Y. (2008). The mechanism selecting the guide strand from small RNA duplexes is different among argonaute proteins. *Plant Cell Physiol.* 49, 493–500. doi: 10.1093/pcp/pcn043
- Turner, M., Yu, O., and Subramanian, S. (2012). Genome organization and characteristics of soybean microRNAs. *BMC Genomics* 13:169. doi: 10.1186/1471-2164-13-169
- Tuskan, G. A., Difazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., et al. (2006). The genome of black cottonwood, *Populus trichocarpa* (Torr. and Gray). *Science* 313, 1596–1604. doi: 10.1126/science.1128691
- Van de Velde, W., Zehirov, G., Sztamari, A., Debreczeny, M., Ishihara, H., Kevei, Z., et al. (2010). Plant peptides govern terminal differentiation of bacteria in symbiosis. *Science* 327, 1122–1126. doi: 10.1126/science.1184057
- Vaucheret, H. (2005). MicroRNA-dependent trans-acting siRNA production. *Sci. STKE* 300, pe43.
- Vaucheret, H. (2008). Plant ARGONAUTES. *Trends Plant Sci.* 13, 350–358. doi: 10.1016/j.tplants.2008.04.007
- Vazquez, F. (2006). *Arabidopsis* endogenous small RNAs: highways and byways. *Trends Plant Sci.* 11, 460–468. doi: 10.1016/j.tplants.2006.07.006
- Vazquez, F., Legrand, S., and Windels, D. (2010). The biosynthetic pathways and biological scopes of plant small RNAs. *Trends Plant Sci.* 15, 337–345. doi: 10.1016/j.tplants.2010.04.001
- Voinnet, O. (2009). Origin, biogenesis, and activity of plant microRNAs. *Cell* 136, 669–687. doi: 10.1016/j.cell.2009.01.046
- Wang, H. W., Noland, C., Siridechadilok, B., Taylor, D. W., Ma, E., Felderer, K., et al. (2009). Structural insights into RNA processing by the human RISC-loading complex. *Nat. Struct. Mol. Biol.* 16, 1148–1153. doi: 10.1038/nsmb.1673
- Wang, X. B., Jovel, J., Udornporn, P., Wang, Y., Wu, Q., Li, W. X., et al. (2011). The 21-nucleotide, but not 22-nucleotide, viral secondary small interfering RNAs direct potent antiviral defense by two cooperative argonautes in *Arabidopsis thaliana*. *Plant Cell* 23, 1625–1638. doi: 10.1105/tpc.110.082305
- Wang, Y., Li, P., Cao, X., Wang, X., Zhang, A., and Li, X. (2009). Identification and expression analysis of miRNAs from nitrogen-fixing soybean nodules. *Biochem. Biophys. Res. Commun.* 378, 799–803. doi: 10.1016/j.bbrc.2008.11.140
- Wong, C. E., Zhao, Y. T., Wang, X. J., Croft, L., Wang, Z. H., Haerizadeh, F., et al. (2011). MicroRNAs in the shoot apical meristem of soybean. *J. Exp. Bot.* 62, 2495–2506. doi: 10.1093/jxb/erq437
- Yan, J., Cai, X., Luo, J., Sato, S., Jiang, Q., Yang, J., et al. (2010). The reduced leaflet genes encode key components of the trans-acting small interfering RNA pathway and regulate compound leaf and flower development in *Lotus japonicus*. *Plant Physiol.* 152, 797–807. doi: 10.1104/pp.109.140947
- Young, N. D., Debelle, F., Oldroyd, G. E., Geurts, R., Cannon, S. B., Udvardi, M. K., et al. (2011). The *Medicago* genome provides insight into the evolution of rhizobial symbioses. *Nature* 480, 520–524
- Zhai, J., Jeong, D. H., De Paoli, E., Park, S., Rosen, B. D., Li, Y., et al. (2011). MicroRNAs as master regulators of the plant NB-LRR defense gene family via the production of phased, trans-acting siRNAs. *Genes Dev.* 25, 2540–2553. doi: 10.1101/gad.177527.111
- Zhang, X., Xia, J., Lii, Y. E., Barrera-Figueroa, B. E., Zhou, X., Gao, S., et al. (2012). Genome-wide analysis of plant nat-siRNAs reveals insights into their distribution, biogenesis and function. *Genome Biol.* 13, R20. doi: 10.1186/gb-2012-13-3-r20
- Zhou, Z. S., Zeng, H. Q., Liu, Z. P., and Yang, Z. M. (2012). Genome-wide identification of *Medicago truncatula* microRNAs and their targets reveals their differential regulation by heavy metal. *Plant Cell Environ.* 35, 86–99. doi: 10.1111/j.1365-3040.2011.02418.x
- Zhu, H., Hu, F., Wang, R., Zhou, X., Sze, S. H., Liou, L. W., et al. (2011). *Arabidopsis* Argonaute10 specifically sequesters miR166/165 to regulate shoot apical meristem development. *Cell* 145, 242–256. doi: 10.1016/j.cell.2011.03.024

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 14 February 2013; accepted: 15 June 2013; published online: 10 July 2013.

Citation: Bustos-Sanmamed P, Bazin J, Hartmann C, Crespi M and Lelandais-Brière C (2013) Small RNA pathways and diversity in model legumes: lessons from genomics. *Front. Plant Sci.* 4:236. doi: 10.3389/fpls.2013.00236

This article was submitted to *Frontiers in Plant Genetics and Genomics*, a specialty of *Frontiers in Plant Science*.

Copyright © 2013 Bustos-Sanmamed, Bazin, Hartmann, Crespi and Lelandais-Brière. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.