

## Bioinformatic characterization and expression analysis of miRNAs in *Solanum lycopersicum*

Ehsan Valiollahi<sup>1\*</sup>, Mohammad Farsi<sup>1</sup>, Pedro Fevereiro<sup>2,3\*</sup>, Amin Mirshamsi Kakhki<sup>1</sup>

<sup>1</sup>Department of Biotechnology and Plant Breeding, College of Agriculture Ferdowsi University of Mashhad, Mashhad, Iran

<sup>2</sup>Instituto de Tecnologia Química e Biológica, Laboratório de Biotecnologia de Células Vegetais, Universidade Nova de Lisboa, Apartado 127, 2781-901 Oeiras, Portugal

<sup>3</sup>Departamento de Biologia Vegetal, Faculdade de Ciências da Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal

\*Corresponding author: valiollahie@gmail.com, psalema@itqb.unl.pt

### Abstract

MicroRNAs (miRNAs) are a class of small endogenous non-coding RNAs that modulate gene expression in both plants and animals and are involved in several biological processes, ranging from organ differentiation to biotic and abiotic stress responses. The comparative analysis of the target genes for miRNAs families may augment our understanding of conserved regulatory mechanisms in the plant kingdom. Here, we used a high throughput sequencing and computational approach to identify new miRNAs in *S. lycopersicum*. Using this approach 18 miRNAs were identified in *S. lycopersicum* leaf samples. We found a closely related mature sequence for Sly-miR319 that derived from distinct precursor sequences and genomic loci. Furthermore, a qRT-PCR-based assay was performed to validate the prediction of five identified miRNAs. Identification of miRNA targets showed that most of the identified miRNAs such as Sly-miR156a, Sly-miR164, Sly-miR166a, Sly-miR172a, Sly-miR319a and Sly-miR482c, modulate the expression of transcription factors that regulate plant development, signaling and metabolism. In brief, our reliable computational analysis of miRNAs and their targets suggests that many important miRNA-targets circuits are also conserved in *S. lycopersicum* and fulfill critical roles in growth and development. These findings provide valuable information for functional characterization of miRNAs in *S. lycopersicum*.

**Keywords:** *Solanum lycopersicum*; High-throughput sequencing; MicroRNAs.

**Abbreviations:** miRNA\_MicroRNA; pre-miRNA\_MicroRNA precursor; miRNA\*\_MicroRNA star; RISC\_ RNA-induced silencing complex; TF\_ transcription factors qRT-PCR\_quantitative real-time PCR.

Received: 11 Oct 2013; Revised: 14 Jan 2014; Accepted: 16 Jan 2014.

### Introduction

MicroRNAs (miRNAs) are a class of small endogenous non-coding RNAs that modulate gene expression in both plants and animals (Carrington and Ambros, 2003, Voinnet, 2009). Mature miRNAs are derived from single-stranded RNA transcripts that possess an imperfect stem-loop secondary structure (Jones-Rhoades et al., 2006). These hairpins are processed by DCL1 into the miRNA/ miRNA\* duplex in the nucleus and are transported to the cytoplasm in plants (Reinhart, 2002, Eamens et al., 2011). The miRNAs are incorporated into the RNA induced silencing complex (RISC), which uses them as guides to recognize target complementary mRNAs and negatively regulate their expression by degradation (Llave et al., 2002) or repression of productive translation (Gandikota et al., 2007). However, recent studies have shown that miRNAs can also positively stimulate the translation of a reporter mRNA when bound to a specific miRNA (miR369-3) where the cultured mammalian cells are serum-starved (Vasudevan et al., 2007). Plant miRNAs play vital roles in multiple essential biological processes, such as leaf morphogenesis and polarity (Palatnik et al., 2003), floral organ identity (Aukerman and Sakai, 2003), and stress responses (Sunkar and Zhu, 2004, Sunkar et al., 2007). Moreover, a major category of plant miRNA target genes have been predicted or confirmed to encode

transcription factors or other regulatory proteins (Bartel, 2004). Thus the activation of miRNAs under stresses will lead to the repression of many protein-coding genes and affect a variety of cellular and physiological processes.

In animals, miRNAs interact with their target sequences in the 3'-UTR through imperfect base-pairing, resulting in the arrest of translation (Bartel, 2009), while plant miRNAs generally recognize their target site through perfect or near-perfect complementary, directing cleavage of mRNAs (Rhoades et al., 2002). This suggests that the miRNA-target modules should be conserved in long evolutionary timescales. As expected, earlier studies demonstrated that the targets of conserved miRNAs have a narrower range of functions than the targets of non-conserved miRNAs (Willmann and Poethig, 2007). Therefore, a comparative analysis of the target genes for miRNAs families may augment our understanding of conserved regulatory mechanisms in the plant kingdom. Identification of miRNAs in different plant species has been based on the cloning of small RNA fractions and/or the prediction of miRNA genes in genomic sequences based on sequence similarity (Jones-Rhoades et al., 2006). These approaches have revealed or predicted several hundred miRNA genes in *Arabidopsis*, rice and poplar, for which complete genomes are available. All

the plant miRNAs reported are available in the Sanger miRNA register (Griffiths-Jones et al., 2008). Recently there has been a surge of interest in identifying miRNAs and profiling their expression pattern using deep sequencing of specifically prepared low-molecular weight RNA libraries. In fact, this method has been used for both purposes in diverse plant species (Fahlgren et al., 2007; Wark et al., 2008, Zhu et al., 2008). In this study we proposed to identify and characterize miRNAs that control development in *S.lycopersicum* via sequencing small RNA libraries derived from *S.lycopersicum* seedling. We employed miRDeep-P (Yang and Li, 2011), utilizing a probabilistic model of miRNA biogenesis to analyze the miRNA transcriptome in plants, to identify potential miRNA precursor transcripts, which provides a deeper understanding of their transcription and regulation. We studied expression patterns of the *S.lycopersicum* miRNAs by stem-loop qRT-PCR technique and predicted potential target genes of most known and new miRNAs in *S.lycopersicum*. Most of the target genes were found to be transcription factors which regulate plant growth, development and metabolism. We believe that, this study will help accelerate our understanding of miRNAs regulatory roles in critical biological processes and these results could be useful for further characterizing the miRNA transcriptome in plants.

## Results and discussion

### Identification of potential miRNAs in *S.lycopersicum*

In this study, more than five million unique *S.lycopersicum* small RNAs were introduced to the non-coding RNA database, Rfam (16,219 sequences) with removed miRNAs, to trim and filter contaminated reads from small ribosomal and nuclear RNAs. 48026 sequences that belong to rRNA, tRNA, snRNA and snoRNA were identified by the Bowtie aligner and were not included in the downstream analysis of miRNA identification. The sequence tags varied in length from 18 bp to 33 bp as seen in Figure 1, with the highest abundance around 23 to 24 bp. The average length of the reads is 23 bp. We compared trimmed sequences to annotated *S.lycopersicum* miRNAs and their precursors that were deposited in miRBase (Release 20). We found that there are 13 overlaps, including Sly-miR156a, Sly-miR156c, Sly-miR162, Sly-miR166a, Sly-miR166b, Sly-miR167, Sly-miR169, Sly-miR171a, Sly-miR171b, Sly-miR172b, Sly-miR4376, Sly-miR482c and Sly-miR6024. Furthermore, we saw three distinct precursor sequences for Sly-miR171a (Table 1). This is probably for improving this miRNA performance in gene regulation. Usually, conserved miRNA are present in different plant lineages and are encoded by multiple loci in the genome to attain high expression levels. In addition to the miRNAs already published and annotated in the miRBase another 5 non reported miRNAs in miRBase and their complementary miRNAs\* were identified in our data (Table 2). The first criteria for microRNA identification is the concomitant presence of the mature miRNAs and the respective miRNA star sequences in the datasets (Rajagopalan et al., 2006). miRNAs are thought to be more stable than miRNA\* *in vivo* due to their combination with RISC (Jones-Rhoades et al. 2006). We found more miRNA sequences than miRNA\* in our data, supporting this observation. In *Arabidopsis* miRNA\* species had ~9% as many reads as the mature miRNAs (Rajagopalan et al., 2006) also supporting our observations. Among the identified miRNAs, Sly-miR164, Sly-miR390, Sly-miR396 and Sly-miR408 are newly found families in *S.lycopersicum* that have

not been published in miRBase. According our analysis, we also found a closely related mature sequence for Sly-miR319 that derived from distinct precursor sequences and genomic loci (Table 2). After the identification of potential miRNAs, A+U content of candidate miRNAs were evaluated. The sequences of the pre-miRNA have A+U content ranging from 50.0% to 67.0% (Table 2), which is in agreement with the conception that miRNA precursors and mature miRNAs contain more A+U nucleotides than G+C (Guddeeti et al., 2005). The average length of miRNA precursors in *S.lycopersicum* is about 70nt, and the variation of the size of the identified miRNAs within different families suggests that they may propose unique functions for tuning their target expression. The location of the mature miRNA sequences in the precursors is variable. In the Sly-miRNA390 and Sly-miRNA164 pre-miRNAs, the mature miRNA sequences were located at the 5' end of the precursors, while they were found at the 3' end in the Sly-miRNA319 and Sly-miRNA408 pre-miRNAs. The stem-loop structures of new miRNA precursors are shown in Fig. 2.

### Verification of high-throughput sequencing data by quantitative real-time PCR

To validate the prediction of miRNAs in *S.lycopersicum* and determine miRNA expression profiles in five weeks old *S.lycopersicum* seedlings, we performed a stem-loop RT-PCR and qRT-PCR assay. Each PCR reaction used one universal reverse primer corresponding to the stem-loop region and one miRNA-specific forward primer. By this procedure, we were able to recognize a positive signal of the expected size at about 60 nt, and validate the presence of these miRNAs. In this experiment all of the putative miRNAs were detected in three replicate, without any nonspecific amplification bands (Fig.3).

### Target gene prediction of miRNAs in *S.lycopersicum*

Gaining insight into the miRNA targets will help us to understand the functional importance of miRNAs. Previous research indicated that miRNA target sites are conserved, thus the conserved sequence can serve as a filter to define putative targets (Ioshikhes et al., 2007; Lall et al., 2006). According to information provided in National Center for Biotechnology Information (NCBI) databases, the identified mRNA targets could be separated into several groups. The largest group encodes transcription factors, which are known to be involved mainly in plant growth and developmental patterning. This is probably a general characteristic of plant miRNAs that tends to be complementary to regulatory targets (Qiu et al., 2007). The next group contained targets encoding a range of different proteins implicated in various metabolic processes, while another group is involved in functions such as hormone responses, stress defense and signaling. Interestingly, a miRNA can be complementary to more than one regulatory target (Table 3). This confirms the possibility that these hits are merely the consequence of complementarity to a nucleotide sequence that encodes a critical and conserved protein motif among gene family members (Rhoades et al., 2002). In the present study, we presented global predictions of the targets (Table 3) of all the known and new *S.lycopersicum* miRNAs identified in this work, which should therefore help to guide experimental verification and potentially uncover important miRNA-target circuits in the future research.

**Table 1.** Mature, microRNA star and precursor sequences of *S.lycopersicum* known miRNAs.

Mature miRBase miRNA	Mature sequence	star sequence	precursors sequence
Sly-miR156a	UUGACAGAAG AUAGAGAGCA C	GCUCUCUAUG CUUCUGUCAU CA	UUGACAGAAGAUAGAGAGCACAUAUGAUGAUUGCU AAUUUCAUUCAAUUUGGGCAGCAAAAAGCAUCUCAU UCAUUUGUGCUCUCUAUGCUUCUGUCAUC
Sly-miR156c	UGUUGACAGA AGAUAGAGAG C	GCUCUCUAUU CUUCUGUCAU C	UGUUGACAGAAGAUAGAGAGCACGAAUAAUGAGGUG CUAAUUGGAAGCUGCACCUUAAUUCUUUGUCUCUCU AUUCUUCUGUCAUC
Sly-miR162	GGAGGCAGCG GUUCAUCGAU C	UCGAUAAACC UCUGCAUCCA G	GGAGGCAGCGGUUCAUCGAUCUGUCCCUGAAAAGCG AUAAAUACAGCAACAGGAAGCGAUCGAUAAACCUCU GCAUCCAG
Sly-miR166a	CUCGGACCAG GCUUCAUCC CC	AGAAUGUUGU CUGGUUCGUG A	AGAAUGUUGUCUGGUUCGUGACCAUAUAUCUUUCU CACAUAAAUGAUCUCGGACCAGGCUUCAUUCCCC
Sly-miR166b	UCGGACCAGG CUUCAUUC C	GGAAUGUUGU CUGGCUCGAG G	GGAAUGUUGUCUGGCUCGAGGUCACCAAUAGAUC AUAAUCCCUAUAUAUACCAUGCUUUUCAAAGAAAG AUUUAAUAUUGAAAAGGGGUUAUGAAUCAUGUGUU AGUGUUGUCGGACCAGGCUUCAUUCCCC
Sly-miR167	AAGCUGCCAG CAUGAUCUA	AGAUCAUGUG GCAGCCUAC C	AAGCUGCCAGCAUGAUCUAAACUUUCCUUUUAGUUCA AAUAAUUCGAGGAAAGAUCAGAUAUGUGGCAGC CUUACC
Sly-miR167	AAGCUGCCAG CAUGAUCUA	CAGAUCAUGU GGCAGCAUCA	AAGCUGCCAGCAUGAUCUAAACUUUUCUAUUGCUCCG UACGAGGAAAGAUCAGAUAUGUGGCAGCAUCA
Sly-miR169a	CAGCCAAGGA UGACUUGCCG G	UGGCAAGUCA UUUUUGGCUA C	CAGCCAAGGAUGACUUGCCGGCAAGCUCGAUCAAUUC UGCUCUAUGUCUGGCAAGUCAUUUUUGGCUAC
Sly-miR171a	UGAUUGAGCC GUGCCAAU C	UGUUGGUGCG GUUCAUAGAG	UGUUGGUGCGGUUCAUUGAGAAAGAAGUUGUUGAAU AAGUUUUUGACCCUACUUUUUGAUUGAGCCGUGCCA AUUUC
Sly-miR171a	UGAUUGAGCC GUGCCAAU C	UGUUGGAAUG GCUCAAU A	UGUUGGAAUGGCUCAAUUCAUUGAAAGUUUCCAAAC AUUUUUGGUCAUUUUAAUUUGAUUGAGCCGUGCCAA UAUC
Sly-miR171a	UGAUUGAGCC GUGCCAAU C	UGUUGGUGUG GUUCAAU A	UGUUGGUGUGGUUCAAUUAGAUAAACAUCUCCACGU AAAAUUAUAGAGUUUGUUUUUUUGAUUGAGCCGUGCC AAUUAUC
Sly-miR171b	AAUUGAGCCG UGCCAAU ACG	AUAUUGGUGC GGUCAAU G	AUAUUGGUGCGGUUCAAUUAGAAAGCGCACUUCUUU AUAAUAGAACUUCGUUAAUUAAUUGAGCCGUGCCA AUAUCACG
Sly-miR172a	AGAAUCUUGA UGAUGCUGCA U	GCAGCACCAU UAAGAUUCAC	GCAGCACCAUUAAGAUUCACAUAGAAAAUAUGGAAC GAAAUUUGCUCAAUUUUUCAUUAUUGAGAAUCUUG AUGAUGCUGCAU
Sly-miR4376	UACGCAGGAG AGAUGAUGCU G	AGCAUCAUAC UCCUGCAU U	UACGCAGGAGAGAUGAUGCUGGACGAUAGAGAAAAU GGGAUUUAAUUUAUUGGCCAGCAUCAUACUCCUGCA UAUU
Sly-miR482c	UCUUGCCAAU ACCGCCAAU CC	AGUGGGUGGU GUGGUAAGAU U	AGUGGGUGGUGUGGUAAGAUUCAUUAUUUUUCC AUAAAAUUGAGAUUCUUGCCAAUACCGCCAAUUC
Sly-miR6024	UUUUAGCAAG AGUUGUUUA CCU	AGAAACAACA CUUGC GAA	AGAAACAACACUUGC UUUUUUUAGCAAGAGUUGUUU UACCU

**Functional analyses of target gene**

Delicate molecular mechanisms are required for plants to accomplish physiological and developmental processes, as well as responses to environmental stimuli. Increasing evidence indicates miRNAs are one essential member of these mechanisms (Jones-Rhoades et al., 2006, Chiou, 2007, Mao et al., 2009). Furthermore, plant miRNAs work cooperatively or antagonistically to establish a balanced regulation. The biological function of miRNAs in *S.lycopersicum* is largely unknown. To present, the function of miRNAs are mainly concluded from production of miRNA-overexpressing transgenic plants or gain of function

mutants in which miRNA-resistant target genes are ectopically expressed. The well-conserved *S.lycopersicum* miRNAs have homologous target genes and operated analogous molecular functions among diverse species in plant kingdom. For instance, miR156 is generally known to target of SBP family plant-specific transcription factor genes, which are involved in early flower development and vegetative phase changes (Wu and Poethig, 2006). The biological function of miR156 was investigated by overexpression in *S.lycopersicum*. Transgenic *S.lycopersicum* plants overexpressing sly-miR156a exhibited a drastically altered phenotype, such as dwarfism, a ‘bush-like’ structure, more abundant leaves, shorter plastochron, later flowering,

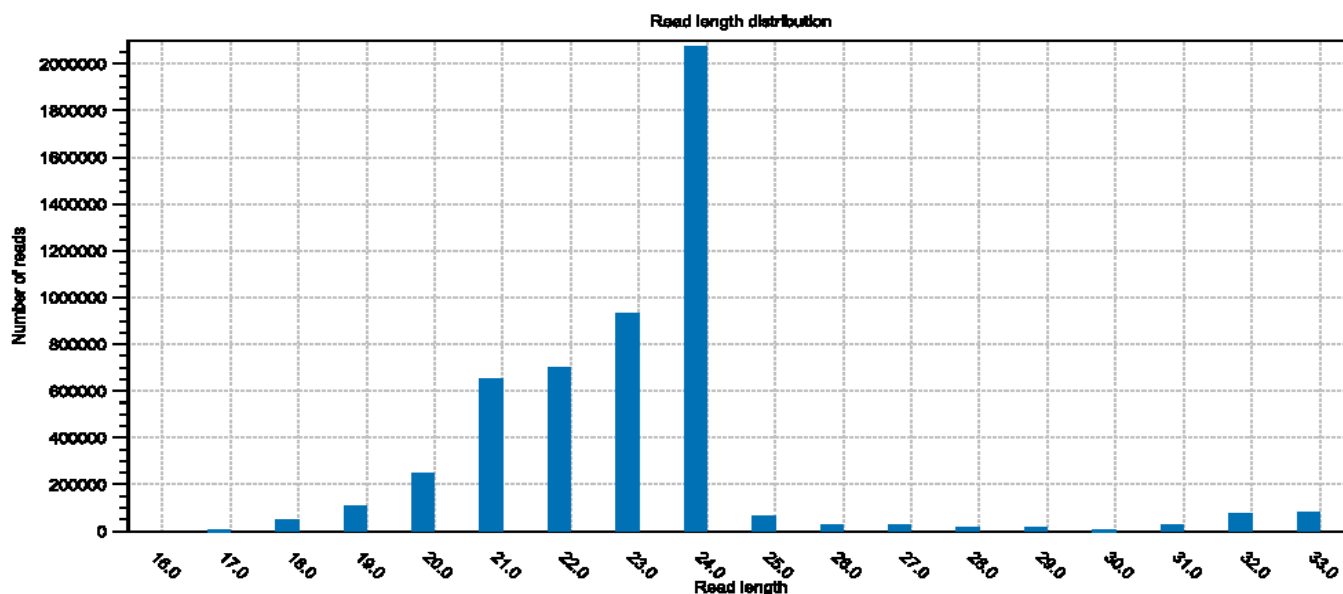


Fig 1. Small RNA read size. Distribution and abundance of small RNA unique sequences in *S.lycopersicum* leaves.

smaller and fewer fruits, and produced numerous adventitious roots (Zhang et al., 2011). The inflorescence structures of miR156 overexpressing plants phenocopied the *sft* mutant (Lifschitz et al., 2006). *S.lycopersicum* sympodial shoot development is regulated by the SFT/SP (Single Flower Truss/Self-Pruning) balance (Shalit et al., 2009), so the aberrant vegetative inflorescence shoots of the transgenic plants may be attributed to the decreased SFT/SP ratio. The targets of miR156 included six (SBP)-box, Transcription factor genes were significantly down-regulated in the miR156-overexpressing plants (Zhang et al., 2011). According to our results, *S.lycopersicum* *No Apical Meristem/Cup-shaped Cotyledon* (NAC) TF genes were identified as targets of miR164. Several NAC-domain genes, including *AtCUC1* and *AtCUC2* are subject to post transcriptional regulation by miR164. Analysis of miR164 mutations and overexpression, as well as of miR164-insensitive *CUC* forms has further revealed the importance of these genes for proper plant development (Mallory et al., 2004; Sieber et al., 2007). As shown in previous research, miR166 in other species targets TF genes of the class III Homeodomain Leucine Zipper (HD-ZIP) family (Sun, 2011). Evidence of targeting of two *S.lycopersicum* HD-ZIP genes was found, including one that had been validated as a target before (Moxon et al., 2008). The phytohormone auxin plays critical roles during plant growth, many of which are mediated by members of the auxin response transcription factor (ARF) family (Guilfoyle et al., 1998). Recent studies have shown that miR167 is complementary to *ARF6* and *ARF8* (Wu et al., 2006). Our results suggest that miR167 and the target *ARFs* are conserved in *S.lycopersicum* (Table 3). Increased levels of auxin accelerate proteolysis of Aux/IAA proteins, which allows ARF proteins to homodimerize and impose their regulatory functions on early auxin-response gene expression (Dharmasiri and Estelle, 2002; Liscum and Reed, 2002). Our prediction shows, the A-subunit genes of the Nuclear Factor Y (NF-Y) or HAP TF and the GRAS/SCR (Scarecrow-related) TF family are targets of *S.lycopersicum* miR169 and miR171, respectively. HAP protein was shown to have an important role during nodule development in *Medicago*

*truncatula* (Comber et al. 2006) and affects flowering time (Wenkel et al. 2006) in *Arabidopsis*. Moxon Simon et al. reported that, miR171 has specific role in early fruit formation. Ethylene-responsive element binding proteins (EREBPs) and APETALA2 (AP2) are prototypic members of a plant-specific family of AP2/EREBP transcription factors (Nole-Wilson and Krizek, 2000; Riechmann and Meyerowitz, 1998; Theissen and Saedler, 1999). Recent studies using *A.thaliana* showed that miR172 and its *AP2-Like* target genes regulate flowering time and floral organ identity (Aukerman and Sakai, 2003; Chen, 2004). Furthermore, *Glossy15*, an *AP2-like* gene from maize that regulates leaf epidermal cell identity, is also detected as a target of miR172 (Lauter et al., 2005). In this research, four predicted targets of miR172 were found to be members of the *AP2* gene family (Table 3). Our identified Sly-miR319a sequence targets *GAMYBlike* gene, one of the important mediators TF in the gibberellin (GA) dependent signaling pathway. Protein phosphatase 2C is key component and repressor of the abscisic acid (ABA) signaling pathway that regulates numerous ABA responses, such as stomatal closure, osmotic water permeability of the plasma membrane (Pos), drought-induced resistance and rhizogenesis genes, response to glucose, high light stress, seed germination and inhibition of vegetative growth (Rodriguez, 1998). According to our prediction, it is found fine-tuned by Sly-miR390 and Cathepsin B-like cysteine proteinase which play an important role in intracellular protein degradation regulated by Sly-miR396. Furthermore Sly-miR482c and Sly-miR6024 were predicted to have largely similar targets among the *CC-NBS-LRR* (coiled coil-nucleotide binding-leucinerich repeat domain) genes that involved in pathogen recognition and biotic stress response.

## Materials and Methods

### Plant materials

Seeds were planted as separate groups in the same pots (20 cm × 12 cm in width and 12 cm in depth) with soil, and incubated in a growth chamber under the following standard

**Table2.** Mature, microRNA star and chromosomal location of precursor sequences for *S.lycopersicum* non reported miRNAs in miRBase.

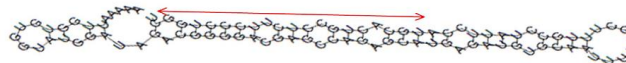
miRNA	mature sequence	star sequence	A+U content	Pre-miRNA sequences	location
Sly-miR319a	UUGGACUGAAG GGAGCUCCCU	GAGCUCUCUC GGUCCAGUC	53.00%	GGAGCUCUCUCCGGUCCAGUCCGAGG CACUCAAGGCCUAAAAACACAGCUGA CGACUCGUUGAUUCUCAACAUAUCA AGAAGCAUUUUUUGAAAUUGGUGUGU UUAAGAUUCAACGAUGCGUGAGCUAU GUUUGGCUAGCGUUGUCACGUCUUGG ACUGAAGGGAGCUCCCU	Ch05:2723775..2723947
Sly-miR390	AAGCUCAGGAG GGAUAGCGCC	CGCUAUCCAUC CUGAGUUUCA	56.00%	AAGCUCAGGAGGGAUAGCGCCAUGGA UGACUCAAUUCGAUUUUUGUACACAU CUUUAGCGCUAUCCAUCUGAGUUUC A	Ch09 8388201..8388279
Sly-miR408	UGCACUGCCUC UUCCUGGCU	ACGGGGACGAG CCAGAGCAUG	50.00%	AAAAAUUGGUGUGGUAUG GAUAGACGGGGACGAGCC AGAGCAUGAGAUGUGCAA UUUUUGCUUUGCCUAUUC CAUGCACUGCCUCUCCU UGGCU	Ch01 82800801..82802000
Sly-miR396	UCCACAGCUU UCUUGAACUU	GUCCAAGAAAG CUGUGGGAAA	67.00%	UCCACAGCUUUCUUGAACUUCUUCU UGCUGAAAUUUGAUCUCUAAAUUGAU AAUUUUGAGAUGAGAUUUGAAGCUAU GAAAGUCCAAGAAAGCUGUGGGAAA	Ch07 2628771..2628873
Sly-miR164	UGGAGAAGCAG GGCACGUGCA	UGCACGUGUUC UCCUUCUCCA	60.00%	UGGAGAAGCAGGGCAGCUGCAAAUUC UUGUAUUCGACAAUAUAUGCAUUACU AAUGGUAUUGCAUGAAUUUUGCACGU GUUCUCCUUCUCCA	Ch01 81732060..817321151



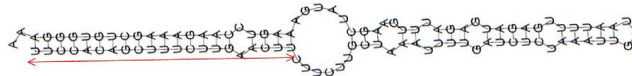
Sly-miR319a



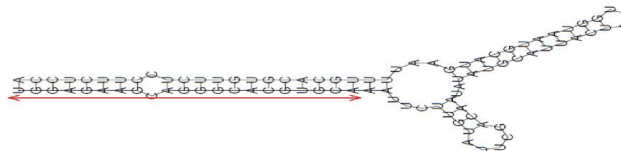
Sly-miR390



Sly-miR408



Sly-miR396

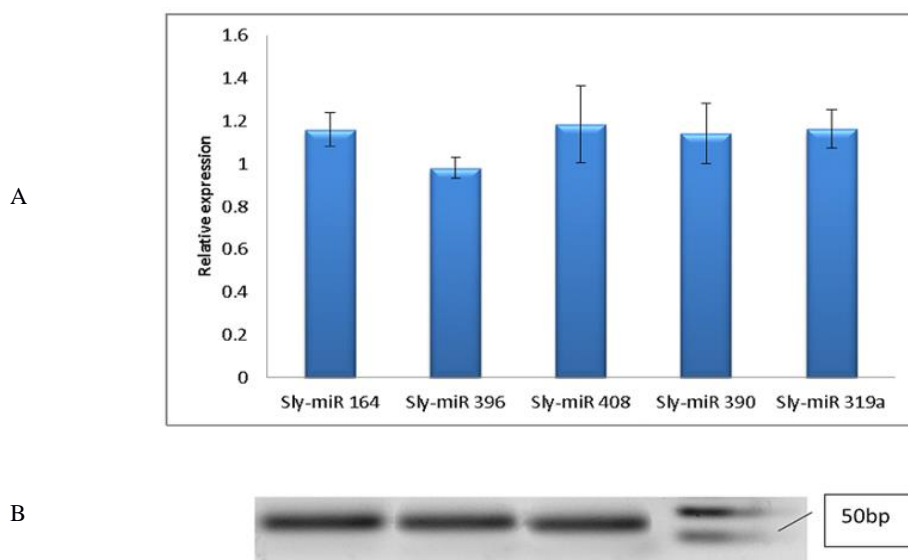


Sly-miR164

**Fig 2.** Stem-loop structures for five identified miRNA precursors, red arrow indicate mature miRNAs sequences

**Table 3.** *S.lycopersicum* miRNAs target prediction.

miRNAs	Target mRNA	ITAG2.3 annotation	Target function
Sly-miR156a	Solyc02g077920.2.1, Solyc04g045560.2.1, Solyc05g012040.2.1, Solyc07g062980.2.1	Squamosa promoter binding-like protein	Transcription factor
Sly-miR162	Solyc04g079470.2.1	Serpin	Serine protease inhibitor
Sly-miR164	Solyc07g066330.2.1, Solyc07g062840.2.1, Solyc06g069710.2.1	NAC domain protein	Transcription factor
Sly-miR166a	Solyc12g044410.1.1, Solyc08g066500.2.1	Class III homeodomain-leucine zipper	Transcription factor
Sly-miR167	Solyc02g037530.2.1	Auxin response factor	Hormone response
Sly-miR169a	Solyc01g006930.2.1, Solyc01g087240.2.1	Nuclear transcription factor Y	Transcription factor
Sly-miR171a	Solyc01g090950.2.1, Solyc08g078800.1.1	GRAS family transcription factor	Transcription factor
Sly-miR172a	Solyc11g072600.1.1, Solyc10g084340.1.1, Solyc09g007260.2.1, Solyc06g075510.2.1	AP2-like ethylene-responsive transcription factor	Transcription factor
Sly-miR319a	Solyc01g009070.2.1	GAMYB	Transcription factor
Sly-miR390	Solyc04g079120.2.1	Protein phosphatase 2C-related	Transcription factor
Sly-miR396	Solyc12g088670.1.1	Cathepsin B-like cysteine proteinase	proteolysis
Sly-miR408	Solyc01g104400.2.1, Solyc04g074740.2.1	Blue copper protein	Metabolic process
Sly-miR4376	Solyc12g006420.1.1	Unknown	Unknown
Sly-miR6024	Solyc11g006640.1.1	Cc-nbs-lrr, resistance protein	Pathogen response
Sly-miR482c	Solyc07g049700.1.1	Cc-nbs-lrr, resistance protein	Pathogen response

**Fig 3.** Expression analysis of candidate miRNAs in leaf tissue by Quantitative RT-PCR (A). Amplification products separated by gel electrophoresis on 2% agarose gel(B).

temperature and light conditions; a photoperiod of 16/8 h day/night, and a thermoperiod of 24/20°C day/night.

#### Construction and sequencing of sRNA libraries

RNA was extracted from the leaves of five week old *S.lycopersicum* seedlings. Small RNA fraction between 19–24 nt was isolated from 15% denaturing polyacrylamide gel and 15 µg was ligated to adaptors without de-phosphorylating and re-phosphorylating (Rathjen et al., 2009). The short RNAs were converted to DNA by RT-PCR and the DNA was sequenced on Illumina HiSeq2000 instrument.

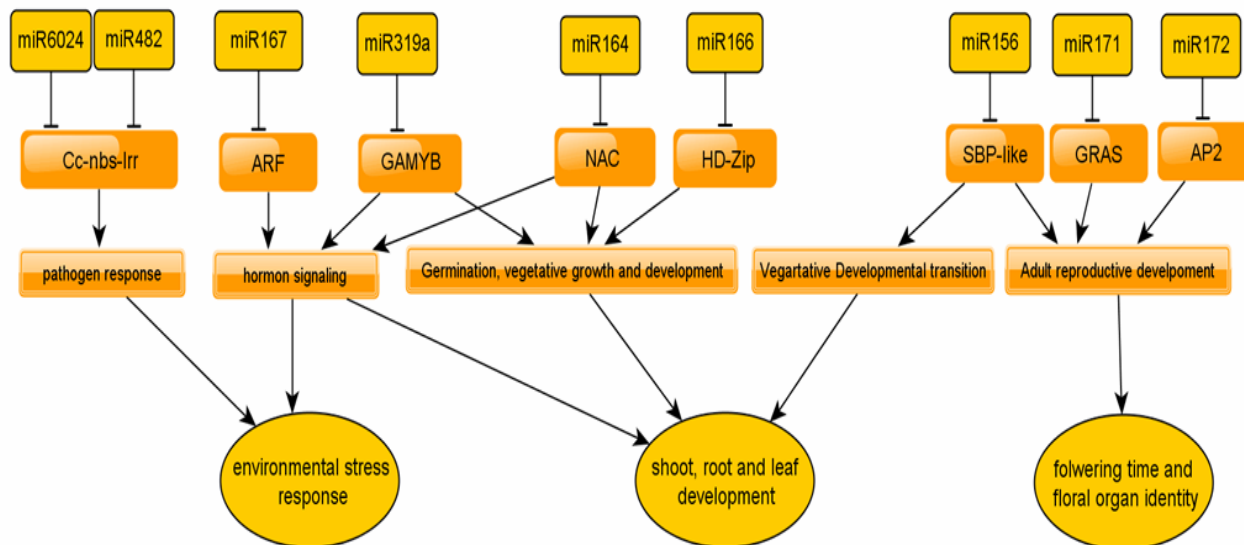
#### Data sources

General information of annotated miRNAs in *S.lycopersicum* was obtained from miRBase (<http://www.mirbase.org/>). Genomic sequences corresponding to the SL2.40 version and

ITAG2.3\_gene\_models.gff3 file were downloaded from The SOL Genomics Network (<http://solgenomics.net/>). Ribosomal and transfer RNAs used to remove non-miRNA reads were collected from Rfam (<http://www.sanger.ac.uk/Software/Rfam/>), the genomic tRNA database (<http://gtrnadb.ucsc.edu/>) and NCBI database ([www.ncbi.nlm.nih.gov/](http://www.ncbi.nlm.nih.gov/)).

#### Identification of miRNAs

More than five million unique reads were obtained from the deeply sequenced small RNA libraries. First, all reads were aligned with the collection of Ribosomal and transfer RNAs via bowtie (allowed two nt mismatches) (Langmead et al., 2009), and unmatched sequences were proceeded to be retrieved for genomic mapping. In details, identical reads from the same sequenced small RNA libraries were collapsed and the copy number recorded. We then utilized Bowtie to



**Fig 4.** Regulatory network of miRNAs in plant growth and developmental processes. miRNAs regulate plant development and stress response via their post-transcriptional control on the expression of many transcription factors and F-box proteins. This control that tunes the regulatory network of plant growth and development is an important level of regulation.

map the collapsed sequence reads into indexed *Solanum lycopersicum* genome. We only kept the reads that were perfectly mapped to the indexed genome (Bowtie parameters  $-a -v0$ ) and parsed them into blast-parsed format. The same strategy was also used to map small RNA reads to the indexed annotated miRNA precursors. The secondary structures of the extracted reference sequences and annotated miRNA precursors, along with all read signature on such sequences were processed by the miRDeep-P algorithms. To identify new miRNAs, 250 bp bracketing sequences based on reads mapped to the Bowtie-indexed genome, were used to be processed with the miRDeep-P package.

#### Target prediction

Target genes of the *S.lycopersicum* miRNAs were predicted by psRNATarget with cDNAs available at the SOL Genomics Network (SGN) database (Dai and Zhao, 2011). Alignments between each miRNA and its putative mRNA target were evaluated based on the following criteria: 1) no more than four mismatches were allowed between the mature miRNA and its target site; 2) no more than one mismatch was allowed at nucleotide positions 1-9; 3) no more than two consecutive mismatches were allowed; 4) no mismatches were allowed at positions 10 and 11 (Xie et al., 2010).

#### Quantitative qPCR expression

After extraction of total RNA from *S.lycopersicum* seedling leaf tissue with TRIzol reagent, RNA was reverse transcribed using Super Script III Reverse Transcriptase (Invitrogen, US) followed by DNase treatment. miRNAs were detected using stemloop RT-PCR method (Varkonyi-Gasic et al., 2007). Stem-loop RT primers were designed according to Chen and coworker (Chen et al., 2005). Primers sequences are reported in Additional File 1. For housekeeping gene, elongation factor, cDNA was reverse transcribed using two specific reverse primers according to the same protocol as for miRNAs. For qPCR analysis, each reaction contained 20 mM Tris-HCl (pH 8.4), 50 mM KCl, 1.5 mM MgCl<sub>2</sub>, 0.1 mM

dNTP mixture, 0.2 μM of each primer: a miRNA specific forward primer and a universal reverse primer (Additional File 1); 1 U of Taq DNA Polymerase Recombinant (Invitrogen), 1 μl of 20X Evagreen dye (Biotium, Inc. Hayward, CA, USA), in a final volume of 20 μl. Real time reactions were performed in a Rotor Gene 6000 (Corbett Mortlake, AUS). For data analysis we employed the “Comparative Quantitation” software supplied by Corbett Research for the Rotorgene. This method for quantitative PCR data analysis calculates the efficiencies of each transcript for each individual PCR reaction and is based on the second differential maximum method to calculate single reaction efficiencies (Rasmussen, 2001).

#### Conclusion

Our reliable computational analysis of miRNAs and their targets suggests that many important miRNA-target circuits are also conserved in *S.lycopersicum*. For instance, as observed in *Arabidopsis* (Wu et al., 2009), both miR156 and miR172, together with their *SBP-like* and *AP2* targets, seem to be involved in the control of phase transition during *S.lycopersicum* development (Figure 4). The conservation of miR164-NAC and miR167-ARF in *S.lycopersicum* indicates that these circuits play potentially important roles in hormone signaling and stress response as was previously validated (Baker et al., 2005; Gutierrez et al., 2009). Our analysis revealed that these conserved miRNA-target circuits could merge into different developmental stages and further coordinate the whole developmental process (Figure 4). The all-important step to understand the biological functions of miRNAs is the search for target genes. In this work, 18 miRNAs were identified from the Illumina sequencing of *S.lycopersicum* leaf samples and their potential targets were predicted, which are related to different physiological processes such as growth and development, stress defense, signaling and metabolic processes. In brief, our results indicate that, the conservation of miRNA-target circuit implies their critical role in *S.lycopersicum* growth and development.

## Acknowledgements

We are grateful to Xiaozeng Yang for informatics assistance. This work was financed by the Ferdowsi University of Mashhad.

## References

- Aukerman MJ, Sakai H (2003) Regulation of flowering time and floral organ identity by a microRNA and its APETALA2-like target genes. *Plant Cell*. 15:2730–2741.
- Baker CC, Sieber P, Wellmer F, Meyerowitz EM (2005) The early extra petals1 mutant uncovers a role for microRNAmiR164c in regulating petal number in *Arabidopsis*. *Curr Biol*. 15(4):303–315.
- Bartel D (2004) MicroRNAs: genomics, biogenesis, mechanism and function. *Cell* 116:281–297.
- Bartel DP (2009) MicroRNAs: target recognition and regulatory functions. *Cell* 136:215–233.
- Carrington JC, Ambros V (2003) Role of microRNAs in plant and animal development. *Science* 301:336–338.
- Chen C, Ridzon DA, Broomer AJ, Zhou Z, Lee DH (2005) Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Res*. 33:179.
- Chiou TJ (2007) The role of microRNAs in sensing nutrient stress. *Plant Cell Environ*. 30:323–332.
- Combiér JP, Frugier F, de Billy F, Boualem A, El-Yahyaoui F, Moreau S, Vernié T, Ott T, Gamas P, Crespi M, Niebel A (2006) MtHAP2-1 is a key transcriptional regulator of symbiotic nodule development regulated by microRNA169 in *Medicago truncatula*. *Genes Dev*. 20(22):3084–8.
- Dai X, Zhao PX (2011) psRNATarget: a plant small RNA target analysis server. *Nucleic Acids Res*. 39:155–159.
- Dharmasiri S, Estelle M (2002) The role of regulated protein degradation in auxin response. *Plant Mol. Biol*. 49:401–409.
- Eamens AL, Agius C, Smith NA, Waterhouse PM, Wang MB (2011) Efficient silencing of endogenous microRNAs using artificial microRNAs in *Arabidopsis thaliana*. *Mol Plant* 4: 157–170.
- Fahlgrén N, Howell MD, Kasschau KD, Chapman EJ, Sullivan CM, Cumbie JS, Givan SA, Law TF, Grant SR, Dangel JL, Carrington JC (2007) High-throughput sequencing of *Arabidopsis* microRNAs: evidence for frequent birth and death of MIRNA genes. *PLoS One*. 2:219.
- Gandikota M, Birkenbihl RP, Höhmann S, Cardon GH, Saedler H, Huijser P (2007) The miRNA156/157 recognition element in the 3'UTR of the *Arabidopsis* SBP box gene SPL3 prevents early flowering by translational inhibition in seedlings. *Plant J*. 49:683–693.
- Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ (2008) osamiR1428aase: tools for microRNA Genomics. *Nucleic Acids Res*. 36:D154–D158.
- Gudeti S, Zhang DC, Li AL, Leseberg CH, Kang H, Li XG, Zhai WX, Johns MA, Mao L (2005) Molecular evolution of the rice miR395 gene family. *Cell Res*. 15:631–638.
- Guilfoyle T, Hagen G, Ulmasov T, Murfett J (1998) How does auxin turn on genes? *Plant Physiol*. 118(2):341–7.
- Gutierrez L, Bussell JD, Pacurar DI, Schwambach J, Pacurar M, Bellini C (2009) Phenotypic plasticity of adventitious rooting in *Arabidopsis* is controlled by complex regulation of AUXIN RESPONSE FACTOR transcripts and microRNA abundance. *Plant Cell*. 21:3119–3132.
- Ioshikhes I, Roy S, Sen CK (2007) Algorithms for mapping of mRNA targets for microRNA. (Review) *DNA Cell Biol*. 26(4):265–272.
- Jones-Rhoades MW, Bartel DP, Bartel B (2006) MicroRNAs and their regulatory roles in plants. *Annu Rev Plant Biol*. 57:19–53.
- Jones-Rhoades MW, Bartel DP, Bartel B (2006) MicroRNAs and their regulatory roles in plants. *Annu Rev Plant Biol*. 57:19–53.
- Lall S, Grun D, Krek A, Chen K, Wang YL, Dewey CN, Sood P, Colombo T, Bray N, Macmenamin P, Kao HL, Gunsalus KC, Pachter L, Piano F, Rajewsky (2006) A genome-wide map of conserved microRNA targets in *C. elegans*. *Curr Biol*. 16:460–71.
- Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*. 10:25.
- Lifschitz E, Eviatar T, Rozman A, Shalit A, Goldshmidt A, Amsellem Z, Alvarez JP, Eshed Y (2006) The tomato FT ortholog triggers systemic signals that regulate growth and flowering and substitute for diverse environmental stimuli. *Proc Natl Acad Sci USA*. 103:6398–6403.
- Liscum E, Reed JW (2002) Genetics of Aux/IAA and ARF action in plant growth and development. *Plant Mol. Biol*. 49:387–400.
- Llave C, Kasschau KD, Rector MA, Carrington JC (2002) Endogenous and silencing associated small RNAs in plants. *Plant Cell*. 14:1605–1619.
- Mallory AC, Dugas DV, Bartel DP, Bartel B (2004) MicroRNA regulation of NAC-domain targets is required for proper formation and separation of adjacent embryonic, vegetative, and floral organs. *Curr Biol*. 14:1035–1046.
- Mao YB, Xue XY, Chen XY (2009) Are small RNAs a big help to plants? *Sci China Ser C-Life Sci*. 52:212–223.
- Moxon S, Jing R, Szittya G, Schwach F, Rusholme Pilcher RL, Moulton V, Dalmay T (2008) Deep sequencing of tomato short RNAs identifies microRNAs targeting genes involved in fruit ripening. *Genome Res*. 18(10):1602–9.
- Palatnik JF, Allen E, Wu X, Schommer C, Schwab R, Carrington JC, Weigel D (2003) Control of leaf morphogenesis by microRNAs. *Nature*. 425: 257–263.
- Qiu CX, Xie FL, Zhu YY, Guo K, Huang SQ, Nie L, Yang ZM (2007) Computational identification of microRNAs and their targets in *Gossypium hirsutum* expressed sequence tags. *Gene*. 395:49 – 61.
- Rajagopalan R, Vaucheret H, Trejo J, Bartel DP (2006) A diverse and evolutionarily fluid set of microRNAs in *Arabidopsis thaliana*. *Genes Dev*. 20:3407–3425.
- Rasmussen R (2001) Quantification on the Light Cycler. In *Rapid Cycle Real-Time PCR: Methods and Applications*. Edited by: Meuer S, Wittwer C, Nakagawara K. New York, NY: Springer-Verlag. 21–34.
- Rathjen T, Pais H, Sweetman D, Moulton V, Munsterberg A, Dalmay T (2009) High throughput sequencing of microRNAs in chicken somites. *FEBS Lett*. 583:1422–1426.
- Reinhart BJ, Weinstein EG, Rhoades MW, Bartel B, Bartel DP (2002) MicroRNAs in plants. *Genes Dev*. 16:1616–1626.
- Rhoades MW, Reinhart BJ, Lim LP, Burge CB, Bartel B, Bartel DP (2002) Prediction of plant microRNA targets. *Cell*. 110:513–520.
- Rodriguez PL (1998) Protein phosphatase 2C (PP2C) function in higher plants. *Plant Mol Biol*. 38(6):919–27.
- Shalit A, Rozman A, Goldshmidt A, Alvarez JP, Bowman JL, Eshed Y, Lifschitz E (2009) The flowering hormone florigen functions as a general systemic regulator of growth and termination. *Proc Natl Acad Sci USA*. 106:8392–8397.



- Sieber P, Wellmer F, Gheyselinck J, Riechmann JL, Meyerowitz EM (2007) Redundancy and specialization among plant microRNAs: role of the MIR164 family in developmental robustness. *Development*. 134:1051–1060.
- Sun G (2011) MicroRNAs and their diverse functions in plants. *Plant Mol Biol*. 18:17–36.
- Sunkar R, Chinnusamy V, Zhu J, Zhu JK (2007) Small RNAs as big players in plant abiotic stress responses and nutrient deprivation. *Trends Plant Sci*. 12:301–309.
- Sunkar R, Zhu JK (2004) Novel and stress-regulated microRNAs and other small RNAs from *Arabidopsis*. *Plant Cell*. 16:2001–2019.
- Varkonyi-Gasic E, Wu R, Wood M, Walton EF, Hellens RP (2007) Protocol: a highly sensitive RT-PCR method for detection and quantification of microRNAs. *Plant Methods*. 3:12.
- Vasudevan S, Tong YC, Steitz JA (2007) Switching from repression to activation: microRNAs can up-regulate translation. *Science*. 318:1931–1934.
- Voinnet O (2009) Origin, biogenesis, and activity of plant microRNAs. *Cell*. 136:669–687.
- Wark AW, Lee HJ, Corn RM (2008) Multiplexed detection methods for profiling microRNA expression in biological samples. *Angew Chem. Int. Ed. Engl*. 47:644–652.
- Wenkel S, Turck F, Singer K, Gissot L, Le Gourrierec J, Samach A, Coupland G (2006) CONSTANS and the CCAAT box binding complex share a functionally important domain and interact to regulate flowering of *Arabidopsis*. *Plant Cell*. 18:2971–2984.
- Willmann MR, Poethig RS (2007) Conservation and evolution of miRNA regulatory programs in plant development. *Curr Opin Plant Biol*. 10:503–511.
- Wu G, Park MY, Conway SR, Wang JW, Weigel D, Poethig RS (2009) The sequential action of miR156 and miR172 regulates developmental timing in *Arabidopsis*. *Cell*. 138:750–759.
- Wu G, Poethig RS (2006) Temporal regulation of shoot development in *Arabidopsis thaliana* by miR156 and its target *SPL3*. *Development*. 133:3539–3547.
- Wu MF, Tian Q, Reed JW (2006) *Arabidopsis* microRNA167 controls patterns of ARF6 and ARF8 expression, and regulates both female and male reproduction. *Development* 133:4211–4218.
- Xie F, Frazier TP, Zhang B (2010) Identification and characterization of microRNAs and their targets in the bioenergy plant switchgrass (*Panicum virgatum*). *Planta*. 232:417–34.
- Yang X, Li L (2011) miRDeep-P: a computational tool for analyzing the microRNA transcriptome in plant. *Bioinformatics*. 27(18):2614–2615.
- Zhang X, Zou Z, Zhang J, Zhang Y, Han Q, Hu T, Xu X, Liu H, Li H, Ye Z (2011) Over-expression of sly-miR156a in tomato results in multiple vegetative and reproductive trait alterations and partial phenocopy of the sft mutant. *FEBS Lett*. 585:435–439.
- Zhu QH, Spriggs A, Matthew L, Fan L, Kennedy G, Gubler F, Helliwell C (2008) A diverse set of microRNAs and microRNA-like small RNAs in developing rice grains. *Genome Res*. 18:1456–1465.