

RESEARCH ARTICLE

Open Access

Identification of novel MiRNAs and MiRNA expression profiling during grain development in indica rice

Ying Lan^{1,2}, Ning Su², Yi Shen¹, Rongzhi Zhang², Fuqing Wu², Zhijun Cheng², Jiulin Wang², Xin Zhang², Xiupin Guo², Cailin Lei², Jie Wang², Ling Jiang¹, Long Mao² and Jianmin Wan^{1,2*}

Abstract

Background: MicroRNAs (miRNAs) modulate gene expression in different tissues and at diverse developmental stages, including grain development in japonica rice. To identify novel miRNAs in indica rice and to study their expression patterns during the entire grain filling process, small RNAs from all stages of grain development were sequenced and their expression patterns were studied using customized miRNA chips.

Results: A total of 21 conserved and 91 non-conserved miRNA families were found in developing indica grains. We also discovered 11 potential novel miRNAs based on the presence of their miRNA*s. Expression patterns of these identified miRNAs were analyzed using customized miRNA chips. The results showed that during the filling phase about half of the detected miRNAs were up-regulated, whereas the remainder were down-regulated. Predicted targets of differentially expressed miRNAs may participate in carbohydrate metabolism, hormone signaling and pathways associated with seed maturity, suggesting potentially important roles in rice grain development.

Conclusions: This study is the first genome-wide investigation of miRNAs during the grain-filling phase of an indica variety of rice. The novel miRNAs identified might be involved in new miRNA regulatory pathways for grain development. The complexity of these miRNAs and their targets and interactions require further study to obtain a better understanding of the molecular mechanisms underlying grain development.

Keywords: miRNA, Grain filling, Indica rice

Background

Small RNAs, microRNAs (miRNAs) and short interfering RNAs (siRNAs) are important gene-regulatory molecules at both the transcriptional and post-transcriptional levels in eukaryotic cells [1]. Plant miRNAs are derived from single RNA molecules. Primary RNA precursors (pri-miRNA) can form imperfect stem-loop structures where a miRNA/miRNA* duplex is processed from the stem by Dicer-like 1 (DCL1) or DCL4 [2-4]. Plant miRNAs negatively regulate their cognate mRNAs by fully or partly binding to complementary sites. After being

methylated at the 3' end by Hua Enhancer 1 (HEN1) [5], the mature miRNA with a length of 20–24 nucleotides (nt) is loaded onto the RNA-induced silencing complex (RISC) to direct the cleavage of its mRNA targets based on extensive complementarity. Plant miRNAs predominantly modulate their targets by mRNA cleavage, and some classes of 24 nt miRNAs direct cytosine DNA methylation at target genes to regulate their expression [6-8]. More recently, miRNA regulation of gene expression *via* DNA methylation and chromatin modification has been suggested [9,10]. The nearly perfect complementarity between miRNAs and their target sites makes it possible to predict their targets by computational approaches. miRNAs were shown to regulate genes involved in basic developmental processes, such as leaf development, flowering time, organ polarity and

* Correspondence: wanjm@njau.edu.cn

¹National Key Laboratory of Crop Genetics and Germplasm Enhancement/Jiangsu Plant Gene Engineering Research Center, Nanjing Agricultural University, Nanjing 210095, China

²National Key Facility for Crop Gene Resources and Genetic Improvement, Institute of Crop Sciences, Chinese Academy of Agriculture Sciences, Beijing 100081, China

auxin signaling [11,12], as well as stress responses and disease resistance [13,14].

High-throughput sequencing technologies allow the discovery of a large set of diverse plant miRNAs. Thousands of miRNAs have been identified in different plant species, rapidly enlarging the identified plant miRNA pool, including miRNAs from different tissues or developmental stages. Based on the recent version of miRBase (<http://www.mirbase.org/>), over 400 miRNAs have been identified in rice. Among them, 21 miRNA families are evolutionarily conserved between Arabidopsis and rice [15-18]. Some of the miRNAs are conserved only among closely related monocots, suggesting the emergence of novel miRNAs after divergence of monocots and dicots [19,20].

As one of the most important food sources for the world's population, rice is also an ideal model plant representing cereal crops. The grain-filling phase is a major stage of plant development that largely determines yield and quality [21]. During this process, all resources of the plant contribute toward a steady rate of starch accumulation in the storage units of rice grains [22,23]. In general, the grain development process can be divided into early development and filling phases. The former is characterized by high biosynthetic activity in grain formation when the total dry matter starts to increase and endosperm starch begins to accumulate rapidly in the seed (6–17 days after flowering, or DAF), whereas during the latter phase (from 18 DAF) the grain usually exhibits a slower increase in dry weight until maximum values are reached and grain weight becomes constant. Global gene expression profiling studies of mRNAs have shown that many genes in multiple pathways participate in grain filling processes, such as those involved in nutrient synthesis, starch synthesis and transport [24,25].

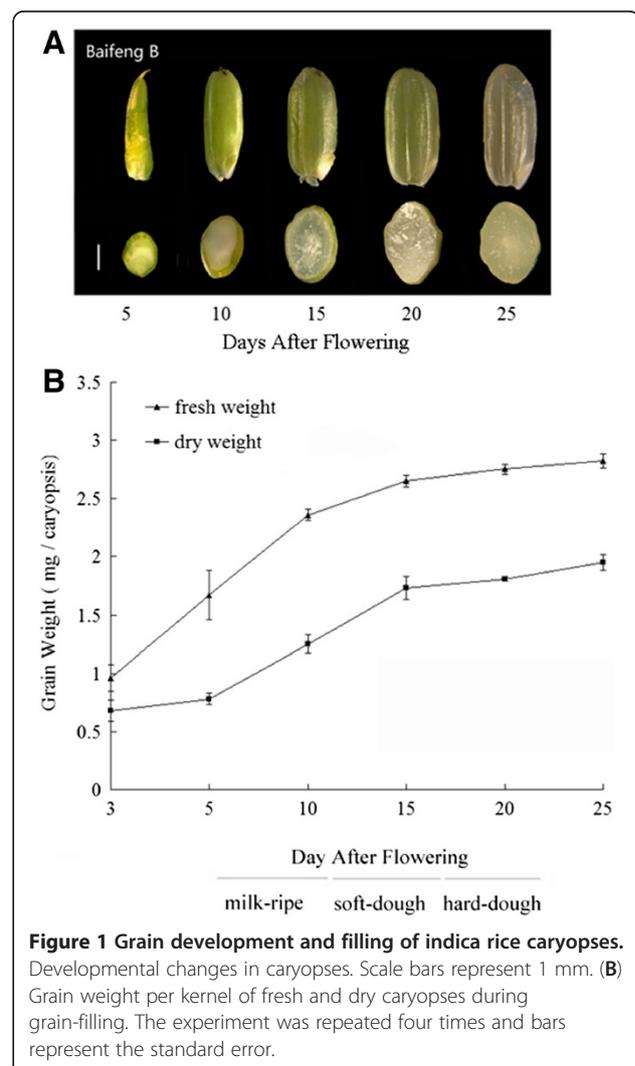
On the other hand, miRNAs were identified as preferentially expressed in various rice organs, including leaf, root, panicle and stem, as well as in seedlings under various stress treatments [17,26,27]. A number of studies were also carried out on small RNAs in the grains of japonica varieties [26,27]. Some miRNAs were preferentially expressed in early developing rice grains, such as 1–10 DAF and 3–12 DAF [28,29], suggesting regulatory roles of miRNAs during grain development. These studies, mainly in subspecies of japonica, also identified significant numbers of both conserved and non-conserved miRNAs. We report here the generation and sequencing of a small RNA library from grain tissues sampled during the entire grain filling stage of an indica cultivar. In addition to numerous conserved miRNAs, we identified 11 novel miRNAs. Subsequently, a customized miRNA chip was generated and miRNA expression profiling was studied using RNA samples from grains of each of the three filling stages: viz. milk-ripe (6–12 DAF), soft-

dough (13–17 DAF), and hard-dough (18–20 DAF). Our results showed that most of the widely conserved miRNAs were down-regulated during grain development whereas rice or grass-specific miRNAs were up-regulated. The targets of differentially expressed miRNAs appeared to be involved in multiple biological processes, such as carbohydrate metabolism, hormone signaling and pathways associated with seed maturity, suggesting that rice miRNAs may play important roles during grain development.

Results

Small RNA populations at the grain filling stage

We measured the fresh and dry grain weights of rice cultivar, Baifeng B, an indica landrace, at several stages of grain-filling (3, 5, 10, 15, 20, and 25 DAF). The fresh weights began to increase from 3 DAF; dry matter accumulation became faster from 5 DAF and reached highest levels at about 25 DAF (Figure 1). Morphological



observations of developing rice seeds showed that the filling phase can be divided into three continuous filling stages. For Illumina sequencing, we isolated small RNAs from immature rice grains sampled at 5 DAF to 25 DAF. After removing low-quality reads, a total of 1,832,288 clean reads were obtained with 974,934 unique sequences. About 637,362 distinct reads (65.4%) were aligned to the '9311' genome using short oligonucleotide alignment program (SOAP) [30] (Additional file 1). Among them, 21 nt and 24 nt small RNAs form the two largest groups, accounting for 22.3% and 50.5% of raw reads, respectively (Additional file 2). By comparison with miRNAs from miRBase v16.0, 102 known miRNAs (71.8%) were found in our dataset. All of these miRNAs, except for miR827, were members of 21 families that are conserved in diverse plant species (Additional file 3A). The abundance of miRNAs varied greatly. MiRNA families highly conserved across plant species, such as miR166, miR167, and miR168, were sequenced more than 10,000 times, whereas previously known stress-induced members, such as miR395 and miR399, were detected less than 10 times (Additional file 3A), indicating that tissue-specific expression patterns of miRNAs are related to their functions. In contrast, most rice- or monocot-specific miRNAs were detected with low read numbers, except for miR444 and miR528, which were represented by 3,917 and 6,305 copies, respectively.

There were significant variations in expression levels for members of the same family. For example, the abundance of the miR159 family varied from 9 (miR159c) to 7,113 (miR159a.2) reads. Similarly, the abundance of members of the miR166 (from 1 to 14,397 reads) and miR164 (from 12 to 6,871 reads) families were also highly variable (Additional file 3A). Twenty previously reported non-conserved miRNA families were not detected in our dataset. A major reason for this might be the limited low sequencing depth, at which the expression level of this group of miRNAs might have been

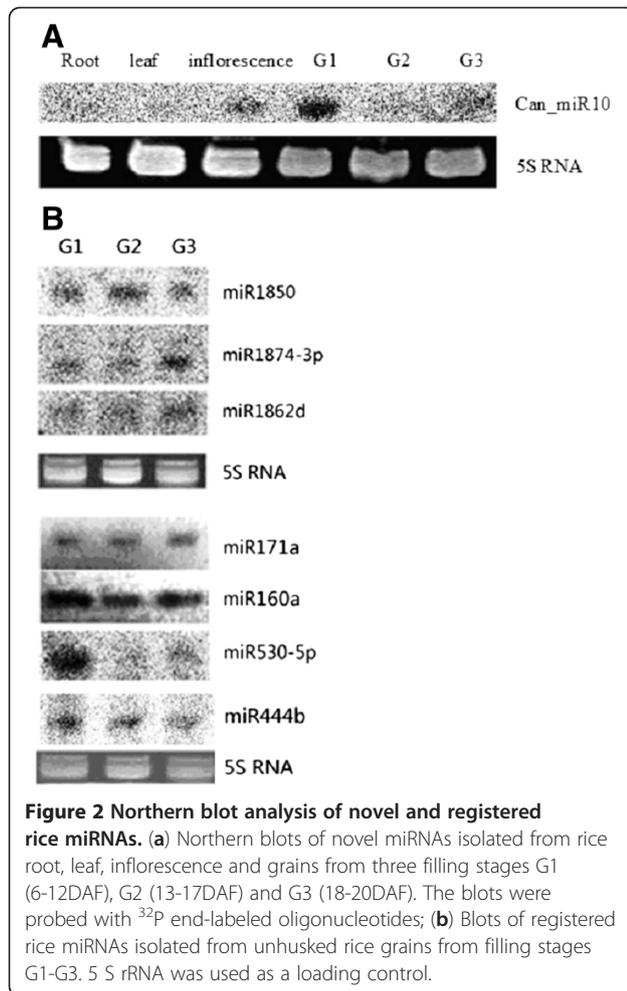
too low to be detected in our library. Another factor may have been the different subspecies and cultivar used compared with previous work. We found that the locations of many miRNA reads varied within a ± 2 nt range from the 5' or 3' ends of annotated miRNA sequences. Some of these variants even had similar reads compared with those annotated in miRBase. For example, the annotated miR1870 had 11 reads in our libraries, whereas the other 22 nt variants had 14 reads. Interestingly, some miRNA*s had higher read numbers than the corresponding miRNAs. For example, miR529* and miR2124* had more reads than their respective miRNAs, 135 vs 0 and 117 vs 1, respectively, suggesting that miRNA* may play a major role in these cases (Additional file 3A).

Identification of 11 novel miRNAs in developing caryopses

To find novel miRNAs, we first mapped all the small RNAs to the sequenced indica cultivar "9311" genome because Baifeng B is an indica landrace. Secondary structures of sequences around the small RNAs were produced using Mfold. These putative miRNA precursors were then used to find miRNA*s, which are considered strong evidence for DCL1-derived products [31]. We found 11 regions that satisfied these criteria and considered them to be novel miRNA gene candidates (Table 1 and Additional file 4). Most novel miRNAs showed weak expression levels (sequencing frequency < 50). The reads for their miRNA*s were even lower. All of these newly identified miRNAs appeared to be rice-specific and had not been reported in other species. Most novel miRNAs were not detectable by northern blotting, except Can_miR 10 (Figure 2A), but all were confirmed by using more sensitive array analysis (see next section). Surprisingly, novel miRNAs discovered in previous deep sequencing of rice grain small RNAs were rarely present in our dataset. Among 39 novel miRNAs

Table 1 Predicted novel miRNAs

Name	Sequence	Length	Abundance	Chromosome	Location
Can_miR_01	ACGGAAAAUCAUGGCUGCACUUAA	24	26	1	Intergenic
Can_miR_02	AAUCAAGUUAGGAACCAUGCAAGU	24	2	6	3'UTR
Can_miR_03	ACUCUAUAUGAACUAAGAUCG	21	6	8	Intergenic
Can_miR_04	UUGGCUGCAUCCCGUUCUCCUC	22	7	4	Intergenic
Can_miR_05	AGCUGCCGACUCAUUCACCCA	21	30	1	Intergenic
Can_miR_06	UCUCUCUCUCCCUUGAAGGCU	21	3	11	Intergenic
Can_miR_07	AUGAAUGUAGGUAAUGCUAGAAAG	24	3	1	Intergenic
Can_miR_08	GAAUGAUCAAAGUUGGACACGAA	23	8	1	Intron
Can_miR_09	ACCUCAACAUGGUUUCAGAACUGG	24	8	7	Intron
Can_miR_10	UCAAUAGCGAUCAAGGCGGAC	21	3	7	Intergenic
Can_miR_11	UGGGAAGGACCAUAAUACCCCUA	24	7	3	Intron



in the 1–10 DAF rice grain library [28], and another 26 reported by Xue et al. [29] in a 3–12 DAF rice grain library, only nine were detected in our library. These included miR1862d and miR1862e with relatively abundant expression levels of 181 and 122 reads, respectively, whereas the others (miR1849, miR1850, miR1856, miR1860, miR1861e, miR2098-3p and miR827b) were detected with expression levels of only one to five reads (Additional file 3B and 3 C). The lack of shared novel miRNAs could be: 1) due to our using indica cultivar Baifeng B whereas all previous studies were with subspecies japonica; 2) because the majority of the rice-specific miRNAs are expressed at very low levels, they might not have been detected at our sequencing depth.

Targets of novel miRNAs were predicted and some appeared to be involved in the grain filling process (Additional file 5). For example, Can_miR_07 was predicted to target starch synthase II, which is preferentially expressed in the endosperm at the middle to later stages of grain filling and plays an important role in elongation of α -1,4 amylase chains [32,33]. Can_miR_04 and Can_miR_08 may target a ubiquitin-protein ligase gene

and a carboxylate oxidase gene, which are known to be involved in cell death and fruit ripening progress [34], respectively. One target of Can_miR_06 is the growth-regulating factor gene (Os04g51190), which is also targeted by miR396 [35], indicating that multiple miRNAs may regulate the same gene family.

MiRNA profile changes during grain filling

To study the expression patterns of miRNAs during grain development, we generated miRNA chips containing 546 probes, and comprising 254 known miRNAs from miRBase version 13.0, the 11 newly identified candidates, and 50 controls. Small RNAs isolated from grains at the milk-ripe stage (6–12 DAF, or phase G1), the soft-dough stage (13–17 DAF, or phase G2), and the hard-dough stage (18–20 DAF, or phase G3) were hybridized to the miRNA chips. The raw signal values are provided in Additional file 6. As shown in Figure 3, 190, 168, and 187 miRNAs were detected above background levels (signal intensity >32) in G1, G2, and G3, respectively. Among them, 143 miRNAs were expressed in all three filling stages, whereas 26, 12, and 30 were specifically expressed in G1, G2, and G3, respectively. Most of the phase-specific miRNAs were newly identified, such as Can miR_11, which is expressed at G1 and G2, Can_miR02 and Can_miR03, which are expressed at G2 and G3, and Can_miR04 and Can_miR11*, which are detected only at G3 (Additional file 6C).

Using a relative intensity change of 2-fold or above between consecutive filling stages, the expression patterns of miRNAs were clustered (Table 2). As shown in Figure 4, 13 miRNA families included 18 members that were differentially expressed across the three filling stages. Nine members of seven miRNA families were

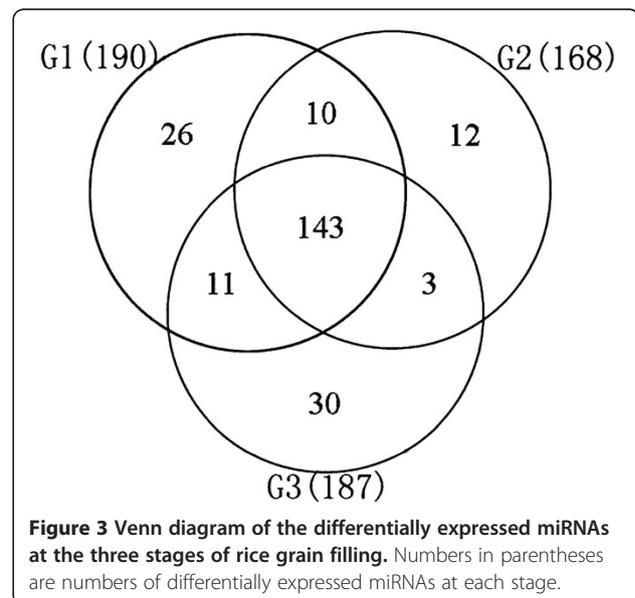


Table 2 Expression data of miRNAs differentially expressed during the three filling stages

Pattern of regulation	miRNA name	MiRNA expression signals (average) ^b			Fold change ^b	Fold change ^b	p-value ^{a,c}
		G1	G2	G3	(G2/G1)	(G3/G2)	
Up	miR1850	56	53	293	0.95	5.53	7.40E-03
	miR1874-5p	349	1140	1266	3.27	1.11	2.29E-05
	miR1862e	133	378	384	2.84	1.02	3.69E-03
	miR1862d	247	526	656	2.13	1.25	3.17E-03
	miR1874-3p	2473	4684	5507	1.89	1.18	3.57E-03
	miR1435	86	138	357	1.60	2.59	2.24E-03
	miR1881	36	65	88	1.81	1.35	1.78E-03
	*miR1867	109	219	186	2.01	0.85	3.14E-03
	miR1884b	99	213	168	2.15	0.79	3.41E-03
Down	miR171a	118	58	42	0.49	0.72	1.13E-03
	miR171b-f	451	253	182	0.56	0.72	4.64E-03
	miR530-5p	4522	3286	1514	0.73	0.46	4.18E-04
	miR160f	125	38	31	0.30	0.82	1.27E-03
	miR160a-d	251	69	38	0.27	0.55	8.01E-04
	miR160e	275	68	48	0.25	0.71	1.16E-03
	miR444b.2	770	443	63	0.58	0.14	3.68E-04
	*miR1873	123	65	47	0.52	0.72	5.21E-05
	miR2055	138	140	97	1.01	0.69	5.23E-03

*: The probe sequence is on the same precursor of this miRNA, but had a ± 2 nt range from the annotated miRNA sequence.

a: p-value < 0.01.

b: Mean levels averaged from three biological replicates; fold changes in miRNA levels were compared from the milky (G1) to hard dough (G3) stages and selected as less than 0.5 or more than 2.0 (from G1 to G3).

c: Signal values <100 (lower than three times the background signal) were not considered for variance analysis.

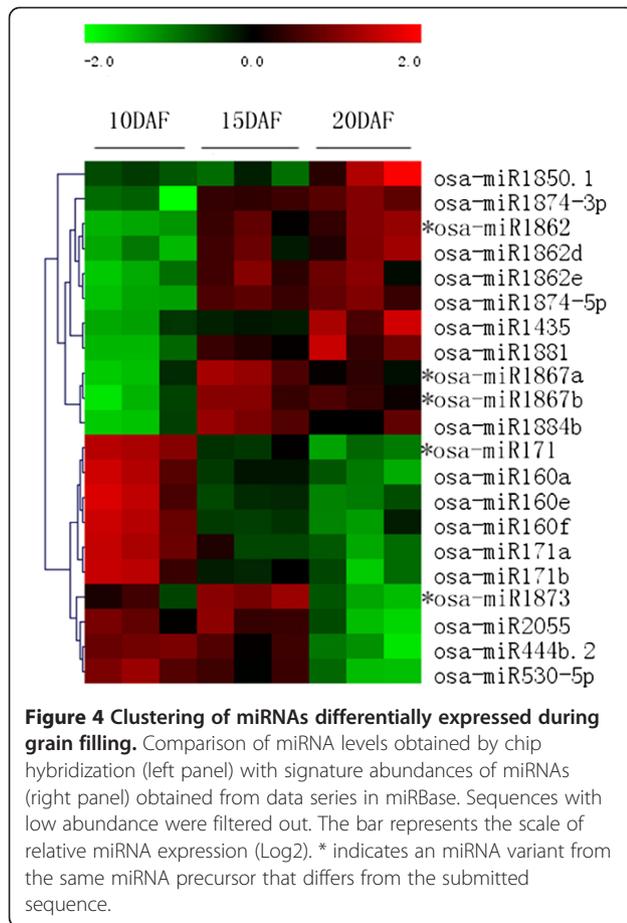
up-regulated. The expression of miR1862 and miR1874 increased from G1 to G2, but remained largely unchanged from G2 to G3, whereas miR159, miR164 and miR1850 underwent rapid increases from G2 to G3. In contrast, nine members of six miRNA families were down-regulated. Among them, the expression of miR160, miR166, and miR171 declined rapidly from G1 to G2, whereas miR167, miR396, miR444 and miR530 gradually declined with advancing grain filling. The expression of miR2055 also declined rapidly from G2 to G3. Some of these expression patterns were consistent with results from northern blot assays. It seems that conserved miRNAs were mostly down-regulated whereas rice- or grass-specific miRNAs were up-regulated during the course of grain filling. As shown in Figure 2B, miR1862, miR1874 and miR1850 were significantly up-regulated, whereas miR171, miR160, miR444 and miR530 were down-regulated. The expression of miR2055 could not be confirmed probably because its expression level was too low.

MiRNA-mediated target mRNA cleavage and target expression patterns during grain filling

To further study the potential effects of differentially expressed miRNAs during grain filling, we computationally

predicted their targets using the miRU program. Rapid amplification of 5' cDNA ends (5' RACE) was used to validate the cleavage events. As shown in Additional file 7A, most targets of conserved rice miRNAs, such as targets of miR160, miR166, miR171, miR444 and miR530, were annotated to be similar to those from other studies. However, the miR1435 target Os04g44354, a UDP-glucuronosyl transferase protein, was not previously reported. Cleavage of Os04g44354 and Os03g43930 occurred with higher frequencies at the 9th and 12th positions of miR1435 and miR166, respectively, in all 12 sequenced clones. This is in contrast to the commonly observed 10th or 11th position of miRNAs, such as the cleavage sites of miR444b.2 on Os04g38780, and miR160 on Os04g43910 and Os04g59430. We also observed a putative target, Os10g30150, for the novel miRNA candidate Can_miR_06, where only three of 10 sequenced clones had cleavage sites at the sixth position; the other degraded fragments were not located on the targeted sequence at all (data not shown).

Finally, quantitative real time-PCR (qRT-PCR) was used to examine the correlation of the expression patterns of miRNAs and their targets. Most of the miRNAs were negatively associated with their targets (Additional file 7B). As shown in Table 3, a large number of targets



of differentially expressed miRNAs during grain filling were transcription factors (Table 3). Based on their biological functions in rice or other species, the predicted target genes appear to be involved in various biological processes. For example, miR159 regulates a MYB gene, which is considered a positive regulator of the GA response during grain maturation [36]. The targets of miR160, Os04g43910 and Os04g59430, are auxin-responsive factors (ARFs), which are important components of auxin signal transduction [37]. MiR444 targeted a type of MADS-box transcription factor that is similar to an Arabidopsis homolog that has roles in fruit dehiscence [38]. Moreover, transcription factors, such as NAC domain proteins, growth factors, and the SCARECROW gene regulator, have been observed in other cellular growth developmental processes. Differential expression of miRNAs and their target genes seem to form a complicated regulatory network that plays a critical role during grain filling in rice.

Discussion

Using high throughput sequencing and customized miRNA chips, we analyzed small RNAs in developing

rice grains from the milky to hard dough stages. The analysis revealed dynamic features of the regulatory network mediated by miRNAs during rice grain development.

Small RNA population and novel miRNAs involved in developing grains

We obtained nearly 2 million high quality small RNAs from grain samples collected from 6 to 20 DAF. A significant proportion of the small RNAs were 21 nt to 24 nt in length. In plants, 21-nt miRNAs and trans-acting siRNAs have roles in post-transcriptional gene silencing by directing mRNA degradation or translational repression [1,39], whereas 24-nt siRNAs tend to be involved in DNA and histone modifications that lead to transcriptional gene silencing [40,41]. Recently, 24-nt miRNAs were also found to direct DNA methylation [42]. In our sequencing data, the reads of 24-nt small RNAs were nearly 7-fold more frequent than reads for 21-nt small RNAs (Additional file 2). The presence of a large population of small RNAs in developing rice grains suggests they have important roles in transcriptional and post-transcriptional regulation of the genes involved in grain development.

Although a number of studies of small RNAs have been carried out using grains from various developmental stages and from various rice accessions, novel miRNAs involved in this process have been continuously discovered [16,26,27]. We sequenced small RNA pools from the developing caryopsis of the indica landrace, Baifeng B, at different stages of development and revealed many classes of conserved miRNAs as well as novel ones. The discovery of 11 novel miRNA candidates was supported by detection of corresponding miRNA*s that were consistent with recent miRNA criteria for characterization [29]. No homologous members were reported in other species, indicating that they are probably rice-specific and found only with extensive tissue sampling.

miRNAs have dynamic expression patterns in developing grains

Many miRNAs display temporal or tissue-specific expression patterns [15,43-45]. Our sequencing results revealed that more than 100 known rice miRNAs were expressed in the rice grain. Several, such as miR156, miR159, miR164, miR166, miR167 and miR396, were expressed at high levels, indicating that, as they are highly expressed in other tissues such as leaf and root, these conserved miRNAs are possibly important regulators for rice plant development.

Our chip data showed that known and novel miRNAs were expressed differentially during the grain filling process. Approximately half of the conserved miRNAs

Table 3 Potential targets of miRNAs with significantly changed expression during rice grain filling

Name	Putative targets (score)	Functions	Pathway
Up-regulated			
miR159	LOC_Os01g12700(1)	Myb-like DNA-binding domain	Signal transduction or/TF
	LOC_Os05g41166(1)	Transcription factor GAMYB	Signal transduction or/TF
miR164	LOC_Os02g36880(1)	NAC domain protein NAC5	Signal transduction or/TF
	LOC_Os06g46270(1)	NAC domain-containing protein 21/22	Signal transduction or/TF
miR1850	LOC_Os04g33510 (2.5)	Expressed protein	Unknown pathway
	LOC_Os09g29200 (3.5)	Glutathione S-transferase	Redox homeostasis
miR1862	LOC_Os02g30730(3)	SART-1 family protein	Cell cycle/biogenesis
	LOC_Os02g48300(4)	Flavonol synthase/flavanone 3- hydroxylase	secondary metabolism
	LOC_Os05g03174(4)	F-box domain containing protein	Protein degradation
	LOC_Os01g53700(4.5)	Glycosyltransferase	Carbohydrate metabolism
	LOC_Os03g58290(4.5)	Mitochondrial-processing peptidase alpha subunit	Amide acid metabolism
	LOC_Os07g22930(4.5)	Granule bound starch synthase GBSS1b,	Carbohydrate metabolism
	LOC_Os07g09460(4.5)	Indole-3-glycerol phosphate lyase	Lipid metabolism
miR1874-3p	LOC_Os01g05680(4)	Geranylgeranyl transferase type-1 beta subunit	Protein folding and transport
	LOC_Os07g38110(5)	Tic20-like protein	Protein folding and transport
miR1874-5p	LOC_Os03g02590(3.5)	Peroxisomal membrane protein PEX11-	Redox homeostasis
	LOC_Os05g18774(3.5)	Protein phosphatase 1	Signal transduction or/TF
	LOC_Os01g16870(4.5)	Argonaute-like protein	RNA processing
	LOC_Os08g28410(4.5)	Anther-specific proline-rich protein APG	Organ development
	LOC_Os06g51084(5)	1,4-alpha-glucan branching enzyme	Carbohydrate metabolism
Down-regulated			
miR160	LOC_Os06g47150(1)	Auxin response factor	Signal transduction or/TF
miR166	LOC_Os03g43930(3.5)	Class III HD-Zip protein 4	Signal transduction or/TF
	LOC_Os03g01890(3.5)	Rolled leaf1	Organ development
miR167	LOC_Os04g57610(3.5)	Auxin response factor 8	Signal transduction or/TF
miR171	LOC_Os02g44360(0)	SCARECROW gene regulator	Signal transduction or/TF
miR396	LOC_Os02g47280(1)	Growth-regulating factor	Cell cycle/biogenesis
miR444	LOC_Os02g36924(0)	MADS-box transcription factor 27	Signal transduction or/TF
	LOC_Os04g38780(0)	MADS-box transcription factor 27	Signal transduction or/TF
	LOC_Os02g49840(1)	MADS-box transcription factor 57	Signal transduction or/TF
	LOC_Os08g06510(2)	Zinc finger, C3HC4 type family protein	Signal transduction or/TF
miR530	LOC_Os02g14990(2)	Zinc finger, C3HC4 type family protein	Signal transduction or/TF
	LOC_Os01g01120(2.5)	E-1 enzyme, putative	Protein degradation
	LOC_Os01g43380(3)	Transferase, transferring glycosyl groups	Carbohydrate metabolism

detected were up-regulated from 6 to 20 DAF, whereas approximately half were down-regulated. Compared with previous reports [28], the expression levels of most miRNAs were approximately the same or up-regulated during the periods 1–5 and 6–10 DAF. Some miRNA genes, such as miR159 and miR399, displayed continually high expression levels throughout grain filling. In contrast, the expression levels of miR160, miR166, miR167, miR171, miR396 and miR444 were down-regulated at the late phase after being up-regulated

during early development (Additional file 8). Previous work on gene expression showed that in the early development phase (6–12 DAF) grain metabolic pathways tend to involve embryo differentiation and cell enlargement [24,46]. This pattern changes at the soft-dough stage (13–17 DAF) and during the late filling phase when grains begin to lose moisture and metabolism switches to senescence and dormancy, processes that might be associated with down-regulated patterns of some miRNAs.

A complex regulatory network in rice grain development

Our results showed that differentially expressed miRNAs seem to regulate large numbers of genes, including many transcription factor genes. In previous microarray analyses, a group of transcription factor genes identified to be involved in the transcriptional control of grain filling included a ZIP type transcription factor that was highly expressed in aleurone and endosperm, and certain MYB genes that may be important in regulating gene expression in developing rice grains [24,47]. On the other hand, NAC domain protein genes regulated by miR164 were implicated in regulating metal mobilization from leaves to seed [48], as well as grain senescence and nutrient remobilization [49], while MADS box transcript genes, the targets of miR444, were considered necessary for fruit ripening in tomato and embryo development in Arabidopsis [16,38,50,51]. In addition, hormonal accumulation and other changes in seeds were shown to affect nitrogen supply and drought tolerance during grain filling [52]; for example, miR160 targets ARFs that can bind auxin-response elements to regulate expression of other genes [37]. Novel miRNAs are often expressed at low levels and match their targets with imperfect pairing. We propose that novel miRNAs may be involved in rice grain development by targeting starch synthesis genes that control the accumulation of starch. Although we were unable to identify the exact cleavage sites on the targets, these novel miRNAs probably regulate their targets by translational inhibition. In light of their important functions in the regulatory network of grain development, future work on these miRNAs and their targets is required.

Conclusions

This work provides the first small RNA expression analysis throughout the entire grain filling phase in an indica rice cultivar. Our small RNA sequencing and chip analysis enlarged the rice miRNA repertoire and confirmed the existence of most conserved, and nearly half of the non-conserved, rice miRNAs in developing grains. Comparison between the three phases of grain filling revealed that these miRNAs and their targets may be involved in diverse pathways, which may also be conserved in other cereal plants.

Methods

Plant materials and construction of a small RNA library

Baifeng B (an indica landrace) was grown under normal field conditions. Immature grains (and husks) were collected at different developmental stages: milk-ripe (5–12 DAF), soft-dough (13–17 DAF) and hard-dough (18–20 DAF). Total RNAs were extracted and equally mixed to construct a library. Small RNAs of 17–27 nt were separated and purified by denaturing polyacrylamide gel

electrophoresis. After dephosphorylation and ligation to an adapter, the products were reverse-transcribed and amplified by PCR, and were later sequenced using Illumina technology.

Bioinformatics analysis and target validation

Primers and 3'/5' adaptors were removed from the original reads and other contaminants (rRNA, tRNA, snRNA and snoRNA) were removed using RepeatMasker. Small RNA sequences of 18 to 26 nt were collected and subjected to BLAST analysis against the *Oryza sativa* ssp. indica 9311 sequence (TIGR Rice Annotation Release 5.0) using SOAP aligner (-s 8 -v 0 -r 2). Whole matching sequences were compared with annotated rice miRNAs and their precursors in miRBase (version 16.0, <http://www.mirbase.org>); homologs of the indica 93–11 genome were regarded as mature miRNAs and miRNA precursors based on PatScan searches (mismatch <= 2, insertion = 0, deletion = 0). MiRNAs located at the position ±2 nt of the precursors were also included as mature miRNAs. New miRNA prediction was based on the rules described by Sunkar et al. [26]. We ran Mfold software using Perl script to identify novel miRNAs; we used a 20 bp frame to search sequences 20 to 260 bp upstream and downstream of each miRNA. Candidate miRNA identification standards were those suggested by Meyers et al. [31], miRNA/miRNA* region with <= 3 bulges, total mismatches <= 6 bases [53]. Candidate targets were identified by miRU following methods previously described [54].

Gene expression analysis using microarray hybridization

Grain samples were collected at three stages: milk-ripe (6–12 DAF), soft-dough (13–17 DAF) and hard-dough (18–20 DAF) with three biological replicates for each stage. Total RNA (5 µg) was used as the starting material for each assay. RNAs were size-fractionated using a YM-100 Microcon centrifugal filter (Millipore, Bedford, MA, USA), and the small RNAs (300 nt) were isolated and extended with a poly(A) tail using poly(A) polymerase. miRNA microarray chips were fabricated by LC Sciences, Houston, Texas, USA. A total of 546 probes were spotted on each chip, including 254 known miRNAs from miRBase version 13.0, 11 newly identified candidates and 50 controls with six duplications. Rice 5 S rRNA served as an inner positive control; and PUC2-20B, an artificial non-homologous nucleic acid, was used as an external positive control. Perfect match and single-base mismatch counterparts to the external positive control, named PUC2PM-20B and PUC2MM-20B, were spiked into the RNA samples before probe labeling. Blank and non-homologous nucleic acids were used as negative controls. Chip hybridization experiments were carried out in triplicate using different

biological samples. Hybridization images were collected using a laser scanner (GenePix 4000B, Molecular Devices, Sunnyvale, CA, USA) and digitized using Array-Pro image analysis software (Media Cybernetics, Silver Spring, MD, USA). Signal values were derived by background subtraction and normalization. A transcript to be listed as detectable had to fulfill at least two conditions: signal intensity higher than $3\times$ (background standard deviation) and spot coefficient of variation (CV) < 0.5 . CV was calculated by (standard deviation)/(signal intensity). When repeating probes were present on an array, a transcript was listed as detectable only if the signals from at least 50% of the repeating probes were above detection level [55]. Student's t-tests were used to assess differences in miRNA levels. To minimize noise and improve accuracy, some probes detected with low abundance (signal value < 100) were not included in variance analysis. Signals below the background average (signal value < 32) were considered non-expressing (Additional file 3).

Northern blot analysis

Low molecular weight RNA (30 μg) was loaded per lane, resolved on a 15% denaturing polyacrylamide gel, and transferred electrophoretically to Hybond-N⁺ membranes (Amersham Biosciences, UK). Both sides of membranes were UV cross-linked for 2 minutes and baked for 1 h at 80°C. DNA oligonucleotides complementary to miRNA sequences were end-labeled with γ -³²P-ATP using T4 polynucleotide kinase (NEB, Beijing). Membranes were hybridized in hybridization buffer (Ambion, US) for 16 h at 42°C. Blots were washed three times with $1\times$ saline sodium citrate and 0.5% sodium dodecyl sulfate at 42°C. Membranes were briefly air-dried and wrapped with Saran Wrap. Images were acquired using a Molecular Imager FX instrument (BioRad, USA).

RNA ligase-mediated 5' RACE and quantitative RT-PCR

Total RNA from rice grain samples that combined equal amounts of material collected at the milk ripe, soft dough and hard dough stages was used to construct a 5'-RACE library. We used the PolyATract mRNA isolation system (Promega) and the GeneRacer kit (Invitrogen) according to the manufacturer's instructions. Two outer and inner-specific primers were used for each RACE reaction (Additional file 9). Amplicons were separated by agarose electrophoresis, cloned into pMD 19-T (Takara) and sequenced. A minimum of six clones were sequenced for each PCR product. In the quantitative RT-PCR experiments of mRNAs, total RNAs were reverse-transcribed using poly(T) adapter. SYBR[®] Green PCR Master Mix (Takara) was used in all quantitative RT-PCR experiments. The relative fold expression changes of target genes were calculated using the 2

delta-delta Ct method. Primers used in all quantitative RT-PCR experiments are listed in Additional file 10.

Additional files

Additional file 1: Data quality of sRNA library.

Additional file 2: Size distribution of small RNAs.

Additional file 3: A The abundance of known miRNAs in the library-1.

Additional file 4: Secondary structures of novel miRNAs.

Additional file 5: Predicted targets of novel miRNAs.

Additional file 6: A Expression patterns of conserved miRNAs.

Additional file 7: Validation and expression of selected miRNA target genes.

Additional file 8: miRNA variant patterns among grain developing stages.

Additional file 9: 5'RLM-RACE primers used in this study.

Additional file 10: Quantitative RT-PCR primers.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

YL, NS and JMW designed the experiment. ZJC and JLW supplied the experimental material. YL and YS collected the samples and carried out the laboratory analyses. YL, RZZ and FQW analyzed the data. YL, LM, and JMW wrote the manuscript. LJ, XPG, JW and JMW supervised the laboratory work. All authors have read and approved the final manuscript.

Acknowledgments

This work was supported by grants from the National Hi-tech Research and Development Program ('863' Program) of China (2012AA100101 and 2007AA10Z137), National Basic Research Program ('973' Program) of China (2009CB118506), and the National Natural Science Foundation of China (31000667).

Received: 13 September 2011 Accepted: 3 April 2012

Published: 21 June 2012

References

1. Voinnet O: Origin, biogenesis, and activity of plant microRNAs. *Cell* 2009, **136**:669–687.
2. Bartel DP: MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 2004, **116**:281–297.
3. Vaucheret H, Vazquez F, Crete P, Bartel DP: The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. *Genes Dev* 2004, **18**:1187–1197.
4. Xie Z, Kasschau KD, Carrington JC: Negative feedback regulation of Dicer-Like1 in Arabidopsis by microRNA-guided mRNA degradation. *Curr Biol* 2003, **13**:784–789.
5. Huang Y, Ji L, Huang Q, Vassilyev DG, Chen X, Ma J-B: Structural insights into mechanisms of the small RNA methyltransferase HEN1. *Nature* 2009, **461**:823–827.
6. Dugas DV, Bartel B: Sucrose induction of Arabidopsis miR398 represses two Cu/Zn superoxide dismutases. *Plant Mol Biol* 2008, **67**:403–417.
7. Lanet E, Delannoy E, Sormani R, Floris M, Brodersen P, Crete P, Voinnet O, Robaglia C: Biochemical evidence for translational repression by Arabidopsis microRNAs. *Plant Cell* 2009, **21**:1762–1768.
8. Wu L, Zhang Q, Zhou H, Ni F, Wu X, Qi Y: Rice MicroRNA effector complexes and targets. *Plant Cell* 2009, **21**:3421–3435.
9. Bao N, Lye KW, Barton MK: MicroRNA binding sites in Arabidopsis class III HD-ZIP mRNAs are required for methylation of the template chromosome. *Dev Cell* 2004, **7**:653–662.
10. Eulalio A, Huntzinger E, Izaurralde E: Getting to the root of miRNA-mediated gene silencing. *Cell* 2008, **132**:9–14.
11. Baulcombe D: RNA silencing in plants. *Nature* 2004, **431**:356–363.

12. Chapman EJ, Carrington JC: **Specialization and evolution of endogenous small RNA pathways.** *Nature Rev Genetics* 2007, **8**:884–896.
13. Chen X: **MicroRNA biogenesis and function in plants.** *FEBS Letters* 2005, **579**:5923–5931.
14. Robert-Seilaniantz A, MacLean D, Jikumaru Y, Hill L, Yamaguchi S, Kamiya Y, Jones JD: **The microRNA miR393 re-directs secondary metabolite biosynthesis away from camalexin and towards glucosinolates.** *Plant J* 2011, **67**:218–231.
15. Jones-Rhoades MW, Bartel DP, Bartel B: **MicroRNAs and their regulatory roles in plants.** *Annu Rev Plant Biol* 2006, **57**:19–53.
16. Lu C, Jeong DH, Kulkarni K, Pillay M, Nobuta K, German R, Thatcher SR, Maher C, Zhang L, Ware D, et al: **Genome-wide analysis for discovery of rice microRNAs reveals natural antisense microRNAs (nat-miRNAs).** *Proc Natl Acad Sci USA* 2008, **105**:4951–4956.
17. Sunkar R: **Cloning and characterization of microRNAs from rice.** *The Plant Cell Online* 2005, **17**:1397–1411.
18. Zhang B, Pan X, Cannon CH, Cobb GP, Anderson TA: **Conservation and divergence of plant microRNA genes.** *Plant J* 2006, **46**:243–259.
19. Fattash I, Voss B, Reski R, Hess WR, Frank W: **Evidence for the rapid expansion of microRNA-mediated regulation in early land plant evolution.** *BMC Plant Biol* 2007, **7**:13.
20. Zhang L, Chia JM, Kumari S, Stein JC, Liu Z, Narechania A, Maher CA, Guill K, McMullen MD, Ware D: **A genome-wide characterization of microRNA genes in maize.** *PLoS Genet* 2009, **5**:e1000716.
21. Mazur B, Krebbers E, Tingey S: **Gene discovery and product development for grain quality traits.** *Science* 1999, **285**:372–375.
22. Anderson A, Hudson M, Chen W, Zhu T: **Identification of nutrient partitioning genes participating in rice grain filling by singular value decomposition (SVD) of genome expression data.** *BMC Genomics* 2003, **4**:26.
23. Krishnan S, Dayanandan P: **Structural and histochemical studies on grain-filling in the caryopsis of rice (*Oryza sativa* L.).** *J Biosci* 2003, **28**:455–469.
24. Zhu T, Budworth P, Chen W, Provart N, Chang HS, Guimil S, Su W, Estes B, Zou G, Wang X: **Transcriptional control of nutrient partitioning during rice grain filling.** *Plant Biotechnol J* 2003, **1**:59–70.
25. Ohdan T, Francisco PB Jr, Sawada T, Hirose T, Terao T, Satoh H, Nakamura Y: **Expression profiling of genes involved in starch synthesis in sink and source organs of rice.** *J Exp Bot* 2005, **56**:3229–3244.
26. Sunkar R, Zhou X, Zheng Y, Zhang W, Zhu J-K: **Identification of novel and candidate miRNAs in rice by high throughput sequencing.** *BMC Plant Biol* 2008, **8**:25.
27. Johnson C, Kasprzewska A, Tennesen K, Fernandes J, Nan GL, Walbot V, Sundaresan V, Vance V, Bowman LH: **Clusters and superclusters of phased small RNAs in the developing inflorescence of rice.** *Genome Res* 2009, **19**:1429–1440.
28. Zhu QH, Spriggs A, Matthew L, Fan L, Kennedy G, Gubler F, Helliwell C: **A diverse set of microRNAs and microRNA-like small RNAs in developing rice grains.** *Genome Res* 2008, **18**:1456–1465.
29. Xue LJ, Zhang JJ, Xue HW: **Characterization and expression profiles of miRNAs in rice seeds.** *Nucleic Acids Res* 2008, **37**:916–930.
30. Li R, Li Y, Kristiansen K, Wang J: **SOAP: short oligonucleotide alignment program.** *Bioinformatics* 2008, **24**:713–714.
31. Meyers BC, Axtell MJ, Bartel B, Bartel DP, Baulcombe D, Bowman JL, Cao X, Carrington JC, Chen X, Green PJ, et al: **Criteria for annotation of plant microRNAs.** *Plant Cell* 2008, **20**:3186–3190.
32. Hirose T, Terao T: **A comprehensive expression analysis of the starch synthase gene family in rice (*Oryza sativa* L.).** *Planta* 2004, **220**:9–16.
33. Nakamura Y, Francisco PB Jr, Hosaka Y, Sato A, Sawada T, Kubo A, Fujita N: **Essential amino acids of starch synthase IIa differentiate amylopectin structure and starch quality between japonica and indica rice varieties.** *Plant Mol Biol* 2005, **58**:213–227.
34. Do YY, Thay TS, Chang TW, Huang PL: **Molecular cloning and characterization of a novel 1-aminocyclopropane-1-carboxylate oxidase gene involved in ripening of banana fruits.** *J Agric Food Chem* 2005, **53**:8239–8247.
35. Sunkar R, Jagadeeswaran G: **In silico identification of conserved microRNAs in a large number of diverse plant species.** *BMC Plant Biol* 2008, **8**:37.
36. Alonso-Peral MM, Li J, Li Y, Allen RS, Schnippenkoetter W, Ohms S, White RG, Millar AA: **The microRNA159-regulated GAMYB-like genes inhibit growth and promote programmed cell death in Arabidopsis.** *Plant Physiol* 2010, **154**:757–771.
37. Ulmasov T, Hagen G, Guilfoyle TJ: **ARF1, a transcription factor that binds to auxin response elements.** *Science* 1997, **276**:1865–1868.
38. Liljegren SJ, Ditta GS, Eshed Y, Savidge B, Bowman JL, Yanofsky MF: **SHATTERPROOF MADS-box genes control seed dispersal in Arabidopsis.** *Nature* 2000, **404**:766–770.
39. Chen X: **Small RNAs and their roles in plant development.** *Annu Rev Cell Dev Biol* 2009, **25**:21–44.
40. Hamilton A, Voinnet O, Chappell L, Baulcombe D: **Two classes of short interfering RNA in RNA silencing.** *EMBO J* 2002, **21**:4671–4679.
41. Zilberman D, Cao X, Jacobsen SE: **ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation.** *Science* 2003, **299**:716–719.
42. Wu L, Zhou H, Zhang Q, Zhang J, Ni F, Liu C, Qi Y: **DNA methylation mediated by a microRNA pathway.** *Mol Cell* 2010, **38**:465–475.
43. Mallory AC, Vaucheret H: **Functions of microRNAs and related small RNAs in plants.** *Nature Genetics* 2006, **38**:S31–S36.
44. Wollmann H, Weigel D: **Small RNAs in flower development.** *Eur J Cell Biol* 2010, **89**:250–257.
45. Rubio-Somoza I, Cuperus JT, Weigel D, Carrington JC: **Regulation and functional specialization of small RNA-target nodes during plant development.** *Curr Opin Plant Biol* 2009, **12**:622–627.
46. Okawa S, Makino A, Mae T: **Effect of irradiance on the partitioning of assimilated carbon during the early phase of grain filling in rice.** *Ann Bot* 2003, **92**:357–364.
47. Onodera Y, Suzuki A, Wu CY, Washida H, Takaiwa F: **A rice functional transcriptional activator, RISBZ1, responsible for endosperm-specific expression of storage protein genes through GCN4 motif.** *J Biol Chem* 2001, **276**:14139–14152.
48. Sperotto RA, Ricachenevsky FK, Duarte GL, Boff T, Lopes KL, Sperb ER, Grusak MA, Fett JP: **Identification of up-regulated genes in flag leaves during rice grain filling and characterization of OsNAC5, a new ABA-dependent transcription factor.** *Planta* 2009, **230**:985–1002.
49. Uauy C, Distelfeld A, Fahima T, Blechl A, Dubcovsky J: **A NAC Gene regulating senescence improves grain protein, zinc, and iron content in wheat.** *Science* 2006, **314**:1298–1301.
50. Busi MV, Bustamante C, D'Angelo C, Hidalgo-Cuevas M, Boggio SB, Valle EM, Zabaleta E: **MADS-box genes expressed during tomato seed and fruit development.** *Plant Mol Biol* 2003, **52**:801–815.
51. Vrebalov J, Ruezinsky D, Padmanabhan V, White R, Medrano D, Drake R, Schuch W, Giovannoni J: **A MADS-box gene necessary for fruit ripening at the tomato ripening-inhibitor (rin) locus.** *Science* 2002, **296**:343–346.
52. Zhang H, Tan G, Yang L, Yang J, Zhang J, Zhao B: **Hormones in the grains and roots in relation to post-anthesis development of inferior and superior spikelets in japonica/indica hybrid rice.** *Plant Physiol Biochem* 2009, **47**:195–204.
53. Jones-Rhoades MW: **Prediction of plant miRNA genes.** *Methods Mol Biol* 2010, **592**:19–30.
54. Zhang Y: **miRU: an automated plant miRNA target prediction server.** *Nucleic Acids Res* 2005, **33**:W701–W704.
55. Thomson JM, Parker JS, Hammond SM: **Microarray analysis of miRNA gene expression.** *Methods Enzymol* 2007, **427**:107–122.

doi:10.1186/1471-2164-13-264

Cite this article as: Lan et al.: Identification of novel MiRNAs and MiRNA expression profiling during grain development in indica rice. *BMC Genomics* 2012 **13**:264.