

RESEARCH ARTICLE

Open Access

# Discovery and comparative profiling of microRNAs in a sweet orange red-flesh mutant and its wild type

Qiang Xu<sup>\*1,2</sup>, Yuanlong Liu<sup>2</sup>, Andan Zhu<sup>2</sup>, Xiaomeng Wu<sup>2</sup>, Junli Ye<sup>2</sup>, Keqin Yu<sup>2</sup>, Wenwu Guo<sup>1,2</sup> and Xiuxin Deng<sup>\*1,2</sup>

## Abstract

**Background:** Red-flesh fruit is absent from common sweet orange varieties, but is more preferred by consumers due to its visual attraction and nutritional properties. Our previous researches on a spontaneous red-flesh mutant revealed that the trait is caused by lycopene accumulation and is regulated by both transcriptional and post-transcriptional mechanisms. However, the knowledge on post-transcriptional regulation of lycopene accumulation in fruits is rather limited so far.

**Results:** We used Illumina sequencing method to identify and quantitatively profile small RNAs on the red-flesh sweet orange mutant and its wild type. We identified 85 known miRNAs belonging to 48 families from sweet orange. Comparative profiling revealed that 51 known miRNAs exhibited significant expression differences between mutant (MT) and wild type (WT). We also identified 12 novel miRNAs by the presence of mature miRNAs and corresponding miRNA\*s in the sRNA libraries. Comparative analysis showed that 9 novel miRNAs are differentially expressed between WT and MT. Target predictions of the 60 differential miRNAs resulted 418 target genes in sweet orange. GO and KEGG annotation revealed that high ranked miRNA-target genes are those implicated in transcription regulation, protein modification and photosynthesis. The expression profiles of target genes involved in carotenogenesis and photosynthesis were further confirmed to be complementary to the profiles of corresponding miRNAs in WT and MT.

**Conclusion:** This study comparatively characterized the miRNAomes between the red-flesh mutant and the wild type, the results lay a foundation for unraveling the miRNA-mediated molecular processes that regulate lycopene accumulation in the sweet orange red-flesh mutant.

## Background

Sweet orange (*Citrus sinensis* [L.] Osbeck), one of the most important fruit crops worldwide, is responsible for 75% of citrus production used both as fresh fruit and processed juice [1]. Orange fruit is believed to contain many phytochemicals which are beneficial to human health [2]. Carotenoids are a group of the health-promoting products which may serve as antioxidants, vitamin A precursors, and cancer-preventing effectors [3]. Orange fruit is reported with more than 115 species of carotenoids [4], but lycopene are absent from common varieties [5-7].

Interest in lycopene metabolism and regulation is growing rapidly because a variety of epidemiological trials have suggested that higher intake of lycopene-containing fruits or foods are associated with decreased cardiovascular disease and prostate cancer risks [8]. So far, three sweet orange mutants with lycopene accumulation in the fruits were reported: Shara [9], Cara Cara [10] and the recently reported 'Hong Anliu' [11]. 'Hong Anliu' is a spontaneous bud mutant from its wild type 'Anliu' sweet orange. We found that lycopene in the 'Hong Anliu' mutant was 1000-fold higher than that in its wild type fruits, and in juice sacs the lycopene accumulation was coincided with increased expression of upstream carotenogenic genes and reduced expression of downstream genes [11]. Thereafter, transcriptional study including cDNA microarray in combination with suppression subtraction hybridization (SSH) was used to investigate gene

\* Correspondence: xuqiang@mail.hzau.edu.cn  
, xxdeng@mail.hzau.edu.cn

<sup>1</sup> Key Laboratory of Horticultural Plant Biology of Ministry of Education, Huazhong Agricultural University, Wuhan 430070, China

<sup>2</sup> Key Laboratory of Horticultural Plant Biology of Ministry of Education, Huazhong Agricultural University, Wuhan 430070, China

Full list of author information is available at the end of the article

expression changes in the mutant, and a total of 267 differentially expressed genes were detected [12]. More recently, massively parallel signature sequencing was applied to decipher the transcriptome changes between WT and MT; and the results indicated that partial impairment of lycopene downstream flux are critical for the formation of lycopene accumulation trait in the mutant [13]. Meanwhile, comparative proteomic analyses were performed between the mutant and its wild type. And the proteomic data, when considered in combination with transcriptional data, suggest that post-transcriptional regulations are involved in shaping the red-flesh trait of the orange mutant [14]. However, it remains largely unknown what kind of post-transcriptional mechanism was involved in the red-flesh mutant. So far, one interesting case has been reported in *Arabidopsis* where miRNA was proposed to regulate the expression of APETALA2 transcription factor (designated as RAP2.2) in controlling a key carotenoids biosynthesis gene (phytoene synthase gene) [15]. But to date, very little is known about miRNAs or miRNA-mediated molecular processes that regulate lycopene accumulation in plant fruits.

Small RNAs (sRNAs) are emerging as important regulators of biological processes on post-transcriptional level in most eukaryotes [16,17]. In plants, sRNAs are involved in a variety of activities that are essential for genome stability, development, and adaptive responses to biotic and abiotic stresses [16-21]. There is still increasing data for the implication of sRNAs in the regulation of other biological processes [22-24].

Based on difference of biogenesis and action, sRNAs have been mainly grouped into two categories, microRNA (miRNA) and short interfering RNA (siRNA) [21]. miRNAs distinguish from siRNAs by their biogenesis. miRNAs are cleaved from stem-loop precursor molecules that derive from single-stranded RNAs, while siRNAs are generated from double-strand RNA precursors [25]. In plants, miRNA regulates gene expression mostly on post-transcription level [17]. The majority of the miRNAs initially isolated from model plant *Arabidopsis thaliana* was found to be evolutionarily conserved across plant species [26]. Later, many researches suggested that the miRNA repertoire of plant species comprises a set of non-conserved miRNAs besides the conserved ones [27-29]. To date, 1894 miRNA sequences have been identified from 24 plant species, as revealed from miRNA registry database (miRBase v13.0; <http://microrna.sanger.ac.uk>). However, no citrus miRNAs have been found in this database so far.

Plant miRNAs have been identified mostly by two strategies: traditional Sanger sequencing of the small RNA libraries as for *Arabidopsis* [19,26], rice [30,31], and poplar [20]; and computational prediction of conserved miRNAs by searching for homologous sequences using EST

or genomic sequences [18,32-36]. However, the computation-based approach is mostly limited to the discovery of conserved miRNAs; and the small-scale traditional sequencing approach identified mainly conserved miRNAs, as suggested by Morin et al. [37] and Moxon et al. [23]. Recently developed high-throughput sequencing technology has allowed the discovery of several non-conserved or lowly expressed miRNAs as well as quantification of the expression of miRNAs. This strategy has been successfully applied to both model plants [28,38-44], and non-model plants [23,37,45-50].

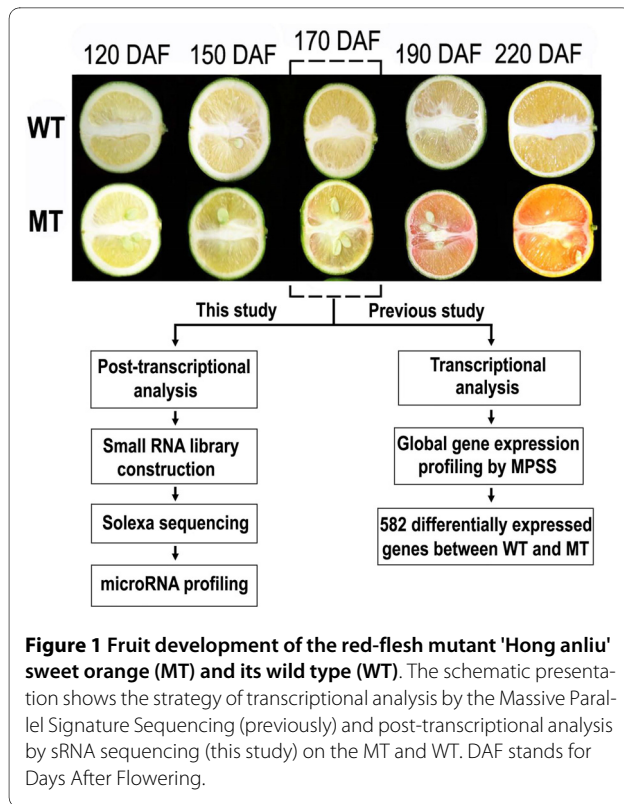
The criteria for miRNAs characterization and annotation in plants have been advanced since plants have relatively large and complex sRNA populations within which miRNAs are often a minority [51]. The minimum criteria include biogenesis criterion that miRNA precursors can be folded into stem-loop structure, and expression criterion that mature miRNAs should be detected by northern blotting, qRT-PCR or sequencing [52]. Proof of biogenesis requires data of the cloning or detection of miRNA\* sequences, which is complementary to miRNA sequences in the precursor molecule [29,51]. Expression detection can also be done by recently developed qRT-PCR-based methods [53,54], where each miRNA is reverse-transcribed from total RNA using a specific stem-loop primer, followed by PCR amplification and detection. This method is believed to be a powerful alternative for sensitive, rapid, and cost-effective detection of miRNA [24,55]. Target prediction can be an ancillary criterion for miRNA annotation. Sequence complementarity between miRNAs and their target genes is very high in plants, which has made the search for plant miRNA target genes a straightforward process [18]. A number of prediction algorithms are available publicly for plant miRNA target prediction, such as miRU [56], WMD3 [57], and UEA plant sRNA toolkit [58].

Here we described the deep sequencing of sRNAs from a red-flesh sweet orange mutant and its wild type. Comparative profiling of the sRNA transcriptome between wild type and mutant revealed a set of 51 known and 9 novel miRNAs showing significant expression changes. The annotation of the potential targets of differential miRNAs indicated that high ranked genes are that implicated in biological processes including carotenoid biosynthesis, transcription regulation, and photosynthesis.

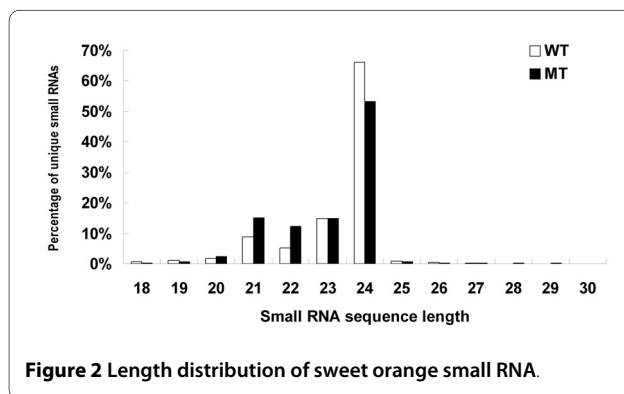
## Results

### Sequencing and annotation of sweet orange small RNAs

Our previous studies showed that the stage of 170 days after flowering (DAF) is the critical stage for the phenotype of color break and global gene expression changes during fruit development in the mutant (Figure 1) [11,13]. In this study, sRNA libraries were generated from fruit pulps of 'Anliu' sweet orange (wild type, WT) and its red-



flesh mutant (MT) 'Hong Anliu' at 170 DAF stage. Deep sequencing of small RNA libraries yielded 4,987,484 and 7,163,041 unfiltered sequence reads for WT and MT, respectively. The raw sRNA data were submitted to Gene Expression Omnibus (GEO) under accession no. GSE18207. After discarding low quality sequences containing ambiguous nucleotides, sequences shorter than 18 nucleotides, and sequences counts below 3 in both libraries, 1,790,395 (WT) and 3,345,746 (MT) reliable sequences were remained for analysis. These sequences represented 202,965 and 320,443 unique sRNA sequences in WT and MT, respectively. The distribution of unique sRNA length is summarized in Figure 2. The overall distribution pattern was similar between WT and



MT, i.e. the majority of sRNAs (approximately 95%) were 21-24 nt in length with 24-nt sRNA as the major peak.

The unique sRNAs were further classified into different RNA categories (Table 1). Analysis showed that 10.8% (WT) and 9.9% (MT) of the unique sRNAs matched the *Citrus sinensis* unigenes (dataset from TIGR, the institute for genomic research). Detailed information was shown in Additional file 1 and 2. Further analysis using the sRNA sequences searching against Rfam database revealed that 1.7%, and 1.5% of the sequences for WT and MT, respectively, matched noncoding RNAs including rRNA, tRNA, snRNA and snoRNAs (Additional file 3 and 4). Searching the remaining sRNAs from WT and MT against miRBase v13.0 identified 0.4% and 0.3% of the unique sequences matched known miRNAs, which account 3.7% and 3.1% of the total sRNAs, respectively (Table 1).

#### Known miRNAs

We identified 85 miRNAs, belonging to 48 families, in sweet orange by BLAST against the miRBase v13.0 using our combined (WT and MT) sRNA sequences (Additional file 5). The potential conserved miRNAs were used for BLASTN against sweet orange unigene dataset to search precursor sequences (Detail in Methods). Since the orange genome is largely unknown, we used orange unigene dataset (from TIGR, the institute for genomic research) and 1.2× coverage sequence of sweet orange (from NCBI) to predict RNA secondary structures. The sequences of perfect match with sRNA sequences were used for fold-back structure prediction by Mfold program [59]. The miRNA-containing fold-back structures were shown in Additional file 6.

To compare the miRNA abundance in different libraries, the count of each miRNA was normalized to transcripts per million (TPM). In the WT sRNA library, miR168, miR172 and miR156 were the most overrepresented families, with the abundance at 8,927, 4,754 and 3,566 TPM, respectively. These three miRNA families also frequently represented in MT sRNA library, but the difference is that the miR172 abundance ranked third (3,330 TPM) after miR156 (4,397 TPM) and miR168 (8,437 TPM) in the library. It is interesting that in both libraries miR168 was the most frequent miRNA which constituted nearly 32% of the miRNA transcriptome of sweet orange.

The 48 miRNA families from sweet orange were compared with those from 11 other plant species (Figure 3). Generally, the orange miRNAs had corresponding homolog in at least one plant species; one miRNA miR1061 only has homolog in phylogenetically distant species *Physcomitrella*. Moreover, miR1169 and miR1171 were extraordinary that they only have homolog genes in *Chlamydomonas* which belongs to protista kingdom.

**Table 1: Summary of sRNAs sequences from WT and MT sweet orange**

Category	Distinct signatures		Total signatures		Mean frequencies	
	WT	MT	WT	MT	WT	MT
matching protein-coding gene						
sense	13811 (6.80%)	18142 (5.66%)	244819 (13.67%)	346569 (10.36%)	17.73	19.10
antisense	8281 (4.08)	10651 (3.32%)	120749 (6.74%)	162839 (4.87%)	14.58	15.29
non-protein-coding RNAs						
snoRNA	13 (0.01%)	14 (0.01%)	471 (0.03%)	149 (0.01%)	36.23	10.64
snRNA	39 (0.02%)	104 (0.03%)	261 (0.01%)	859 (0.03%)	6.69	8.26
tRNA	647 (0.32%)	1301 (0.41%)	21932 (1.22%)	73778 (2.21%)	33.90	56.71
rRNA	2747 (1.35%)	3772 (1.18%)	59535 (3.33%)	99252 (2.97%)	21.67	26.31
miRNAs						
known	667 (0.33%)	802 (0.25%)	61993 (3.46%)	87455 (2.61%)	92.94	109.05
novel	83 (0.04%)	111 (0.03%)	4084 (0.23%)	11463 (0.34%)	49.20	103.27
other sRNAs	176677 (87.05%)	285546 (89.11%)	1276551 (71.30%)	2563382 (76.62%)	7.23	8.98
total	202965 (100%)	320443 (100%)	1790395 (100%)	3345746 (100%)	8.82	10.44

These data suggest that some orange miRNAs are ancient. Sweet orange was phylogenetically close to Arabidopsis which have 187 miRNA genes in miRBase v13.0 database; comparison between them showed that 28 miRNA families were conserved between them. When comparing with the 140 miRNA genes from another fruit crop grapevine, 23 of them were conserved between sweet orange and grapevine (Figure 3).

Putative targets of miRNAs were determined using publicly available algorithms including miRU [56] and WMD3 [57]. Of the 85 known miRNAs, 82 had putative targets, with the exception of miR167a, miR845b, and miR1520c (Additional file 7). Many putative targets are transcription factors and homologs of known miRNA target genes in other plant species, such as SBP for miR156, NAC for miR164, bZIP for miR166, AP2 for miR172 and F-box for miR394. One occasional case in grapevine [60] was also observed in sweet orange: csi-miR535 was predicted to target *squamosa promoter-binding protein-like (SPL)* gene while not target to a brassinosteroid signaling regulator as that firstly reported in moss [61].

#### Novel miRNAs

After excluding the rRNA, tRNA, snRNA, snoRNAs and known miRNAs, the remained sRNA sequences were used to BLAST search on orange unigenes and 1.2× cov-

erage sequence of sweet orange (from NCBI); and secondary structure prediction were done for each locus using the criteria described by Jones-Rhoades et al. [29]. This analysis revealed that 21 sRNA sequences were perfectly matched ESTs and could be folded into stem-loop structures (Additional file 8). Search of complementary miRNA\* sequences showed that 12 of the 21 candidate miRNAs were detected (Table 2 and Additional file 8). Based on the criteria of loci with stem-loop structure producing sequenced mature miRNA and miRNA\* species as described in Meyers et al. [51], these 12 sequences were considered as novel miRNAs.

All the 12 novel miRNAs had a predicted target in at least one orange EST sequence (Additional file 9). Putative targets of four miRNAs were transcription factors, including NAC for csi-novel-01 and csi-novel-11, AUX/IAA TF for csi-novel-03 and YABBY for csi-novel-10. It is notable that one miRNA potentially target disease resistance genes, i.e. csi-novel-06 predicted to target on serine/threonine kinase gene.

To confirm the expression of the novel miRNAs, we analyzed the expression of all the 12 novel miRNAs using stem-loop qRT-PCR. All the genes could be detected by qRT-PCR (Figure 4A). Of the 12 novel miRNAs, 8 genes showed similar expression pattern to that revealed from sequencing data of the sRNA libraries. The electrophore-



**Figure 3** Known miRNAs from sweet orange (*Citrus sinensis*), designated as csi on the first column, and their homologs in other plant species. The picture is modified from Axtell et al. [45]. Scale 0, miRNA only predicted; scale 1, miRNA sequenced; scale 2, miRNA/miRNA\* accumulation detected; scale 3, miRNA detected by small RNA blot or qRT-PCR; scale 4, miRNAs with validated targets. The general information is based on miRBase 13.0 statistics. Abbreviations: ath, *Arabidopsis thaliana*; bna, *Brassica napus*; csi, *Citrus sinensis*; eca, *Eschscholzia californica*; mtr, *Medicago truncatula*; osa, *Oryza sativa*; pta, *Pinus taeda*; ptc, *Populus trichocarpa*; ppt, *Physcomitrella patens*; sly, *Solanum lycopersicum*; vvi, *Vitis vinifera*; zma, *Zea mays*. Some of the data is renewed with information from publications including *Arabidopsis* [24,62], California poppy [46], grape [60], rice [44], and tomato [23,55].

**Table 2: Potential novel miRNA genes from sweet orange**

miRNA ID	Sequences (5'-3')	WT count <sup>a</sup>	MT count <sup>a</sup>	Fold change	Precursor Unigene ID	Start, end	miRNA star
csi-novel-01	UUUUUCGGCAACAU GAUUUCU	2.8	0.0		<a href="#">EY671768</a>	607,764	Yes
csi-novel-02	UGGAGGCAGCGGUU CAUCGAUC	4.5	3.3	1.4	<a href="#">EY681703</a>	286,417	Yes
csi-novel-03	UAGAUAAAGAUGAG AGAAAAA	220.1	1248.5	5.7	<a href="#">EY683731</a>	19,166	Yes
csi-novel-04	UUCAAGAAAUCUGU GGGAAG	8.9	0.0		<a href="#">EY714194</a>	255,397	Yes
csi-novel-05	UGAAGGGCCUUUCU AGAGCAC	90.5	29.6	3.1	<a href="#">CN182895</a>	123,285	Yes
csi-novel-06	UUCCUAGUCCCCCU AUUCCUA	11.2	1.5	7.5	<a href="#">CX051199</a>	168,263	Yes
csi-novel-07	GGAAUGUUGUCUGG CUCGAGG	12.3	4.8	2.6	<a href="#">EY672989</a>	141,318	Yes
csi-novel-08	AGUGGGAGCGUGGG GUAAGAAG	287.1	369.1	1.3	<a href="#">EY676411</a>	137,282	Yes
csi-novel-09	UUGAGUUCUGCAAG CCGUCGA	0.0	2.7		<a href="#">EY712260</a>	398,498	Yes
csi-novel-10	AGGUCAUCUUGCAGC UUCAAU	27.9	4.5	6.2	<a href="#">EY739655</a>	272,634	Yes
csi-novel-11	UGGACAGAGAAAUCA CGGUCA	591.5	132.7	4.5	<a href="#">TC19570</a>	105,191	Yes
csi-novel-12	GCAGGCUCCUCAAGA UUCACA	15.1	4.5	3.4	<a href="#">TC23189</a>	300,444	Yes

<sup>a</sup> count was normalized as transcripts per million (TPM).

sis of the qRT-PCR products confirmed the size and expression of these novel miRNAs (Figure 4C). Moreover, we analyzed the expression of three miRNA\*s (csi-novel-03, -05, -06); and results showed that two miRNA\*s had different expression pattern from those of corresponding mature miRNAs (Figure 4A).

#### miRNA-targets associated small RNAs

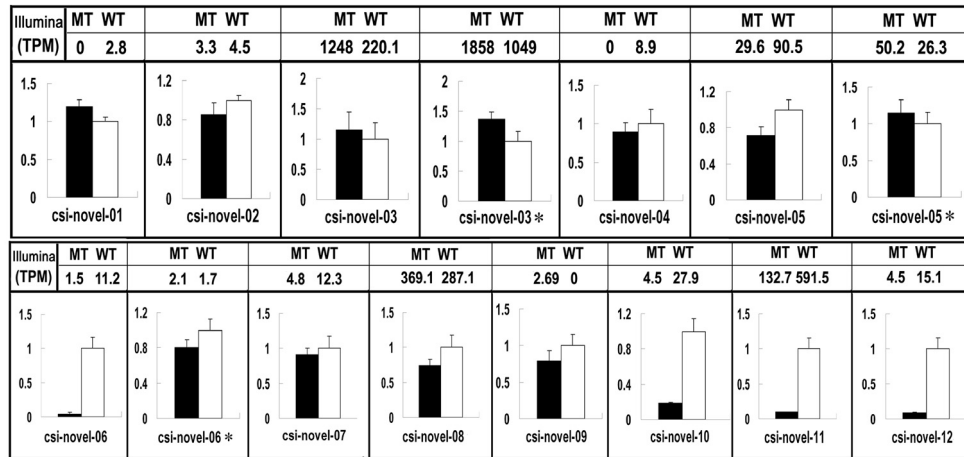
To compare the frequency of small RNAs located in miRNA-target genes with other genes, we mined the unique small RNAs (sRNAs) that matched perfectly to miRNA-target genes in sense or antisense strands according to Luo et al. [62]. Target prediction of the 85 known and 12 novel miRNAs produced a total of 555 miRNA-target genes. Then all the 555 miRNA-target genes were searched against the unigenes on which sRNAs were mapped. This analysis in WT showed that 13,811 sRNAs in sense and 8,281 sRNAs in antisense were mapped perfectly on orange unigenes (Table 1; Additional file 1). Of the mapped unigenes in WT, 76 were miRNA-target genes: 21 of them produced sRNAs in sense strand, 11 in antisense and 44 in both sense and

antisense (Table 3). For MT sample, 18,142 and 10,651 sRNAs were mapped in sense and antisense, respectively (Table 1; Additional file 2). Similar analysis revealed that 97 miRNA-target genes match sRNAs perfectly, with 20 in sense strand, 27 in antisense and 50 in both strand (Table 3; detail in Additional file 10). This study showed that 13.7% and 17.5% of the miRNA-target genes in WT and MT could be the precursors of sRNAs. The high frequency of miRNA-targets associated sRNAs suggests that miRNA-target genes are hotspots for generation of sRNAs.

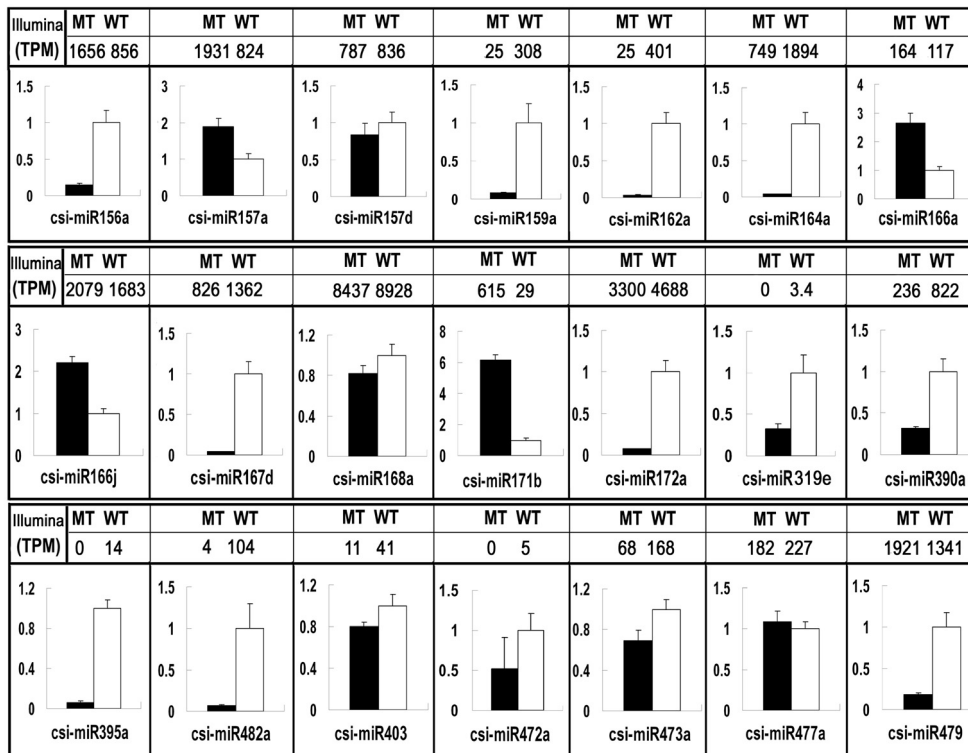
#### Differentially expressed miRNAs between WT and MT

The frequency of each miRNAs in the small RNA library could serve as an index for the estimation of the relative expression abundance of miRNAs. To investigate the expression change of miRNAs between WT and MT, the abundance of each miRNA was normalized as transcripts per million (TPM). A total of 51 known miRNA genes and 9 novel miRNAs were up- or down-regulated significantly at 0.01 level based on Z-score statistics (Table 2 and Additional file 5). Among these 60 differentially

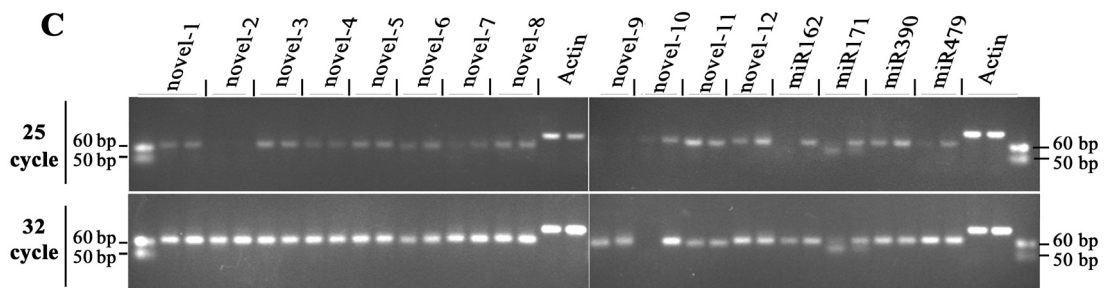
**A**



**B**



**C**



**Figure 4 Expression confirmation of citrus miRNAs derived from high throughput sequencing.** (A). novel miRNAs expression detected by stem-loop qRT-PCR; (B) differentially expressed known miRNAs expression detected by stem-loop qRT-PCR; (C) electrophoresis of the stem-loop PCR products. Each primer has been used for the PCR amplifications on two samples, mutant and wild type.

expressed genes, 52 miRNAs showed a two-fold or greater (ratio > 2 or < 0.5) expression difference between MT and WT. Table 4 summarized the 20 miRNAs with the most significant expression changes between WT and MT.

To confirm the expression difference of the miRNA, we analyzed the expression of 21 known miRNA genes using stem-loop qRT-PCR in addition to the 12 novel miRNA genes described earlier. The expression level of each gene in the MT and WT was compared with its abundance from the sequencing data of the sRNA library. Results showed that among the 33 miRNAs, 18 known miRNAs and 8 novel miRNAs have the same expression pattern in WT and MT as that from sequencing data (Figure 4B).

#### Annotation of potential targets of differentially expressed miRNAs

Functional annotations were further performed to investigate which processes are differentially regulated by miRNAs in the mutant. A total of 418 potential target genes were predicted based on the 60 differentially expressed miRNAs in sweet orange (Additional file 11). Interestingly, the potential targets included two important carotenogenic genes, i.e. geranylgeranyl pyrophosphate synthase (GGPS) gene which targeted by csi-miR167 and lycopene  $\beta$ -cyclase (LYCb) gene which targeted by csi-miR1857. To evaluate the potential functions of these miRNA-target genes, gene ontology (GO) categories were assigned to the putative targets of the 60 differentially expressed miRNA genes according to the method described by Morin et al. [63]. Figure 5 summarized the categorization of miRNA-target genes according to the cellular component, molecular function, and biological process. Based on molecular function, the genes were finally classified into 13 categories, as shown in Figure 5B; the most three over-represented GO terms are metal ion binding, nucleic acid binding, and transcription factor activity. Categories based on biological processes revealed that the miRNA-target genes were related to 17 biological processes; the most three frequent terms are regulation of transcription, signal transduction, protein modification (Figure 5C). Moreover, GO representations from this study were compared with that based on all the

public unigenes from sweet orange in TIGR gene index database [http://compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/GO\\_browser.pl?species=Orange&gi\\_dir=csg](http://compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/GO_browser.pl?species=Orange&gi_dir=csg); the analysis revealed that the striking differences lies in the high percentage of nucleus, plastid and chloroplast for cellular component, and overrepresentation of regulation of transcription, protein modification and photosynthesis for biological process in this study.

The biological interpretation of the target genes of differential miRNAs was further completed using KEGG pathway analysis. A total of 15 different metabolic pathways were found (Table 5), of which some were consistent with biological processes that already revealed by GO analysis. The most represented pathways included carbon fixation (8 enzymes represented), fatty acid metabolism (5), flavonoid biosynthesis (6), terpenoid biosynthesis (3), and ubiquitin mediated proteolysis (3).

High ranked target genes of the differentially expressed miRNAs were those implicated in three biological processes including regulation of transcription, protein modification and photosynthesis. It is easy to understand the high frequency of terms 'regulation of transcription' and 'protein modification' since miRNA are involved in diverse regulatory events [64]. However, the biological process of photosynthesis needs to be verified. Moreover, two carotenogenic genes were also predicted to be targeted by miRNAs. We further analyzed one gene involved in the carbon fixation (miR319a and putative target phosphoenolpyruvate carboxylase [TC11997]) and another gene involved in carotenoid biosynthesis (miR1857 and putative target lycopene  $\beta$ -cyclase [LYCb]). The results showed that the expression profiles of miRNAs are complementary to the profiles of their target genes (Figure 6), implied the possible roles of miRNAs in regulation of the biological processes that involved in shaping the mutant trait in sweet orange.

#### Discussion

We have characterized the small RNA transcriptome in a red-flesh mutant that with abnormal pattern of lycopene accumulation and its wild type. With deep sequencing strategy by Illumina platform, it provides us not only large number of conserved miRNAs but also novel miR-

**Table 3: Summary of small RNAs generated from miRNA-target genes**

	Producing sense small RNAs	Producing antisense small RNAs	Overlap between sense and antisense	Total sense or antisense
WT miRNA target genes <sup>a</sup>	65 (11.7%)	55 (9.9%)	44 (7.9%)	76 (13.7%)
MT miRNA target genes <sup>a</sup>	70 (12.6%)	77 (13.9%)	50 (9.0%)	97 (17.5%)

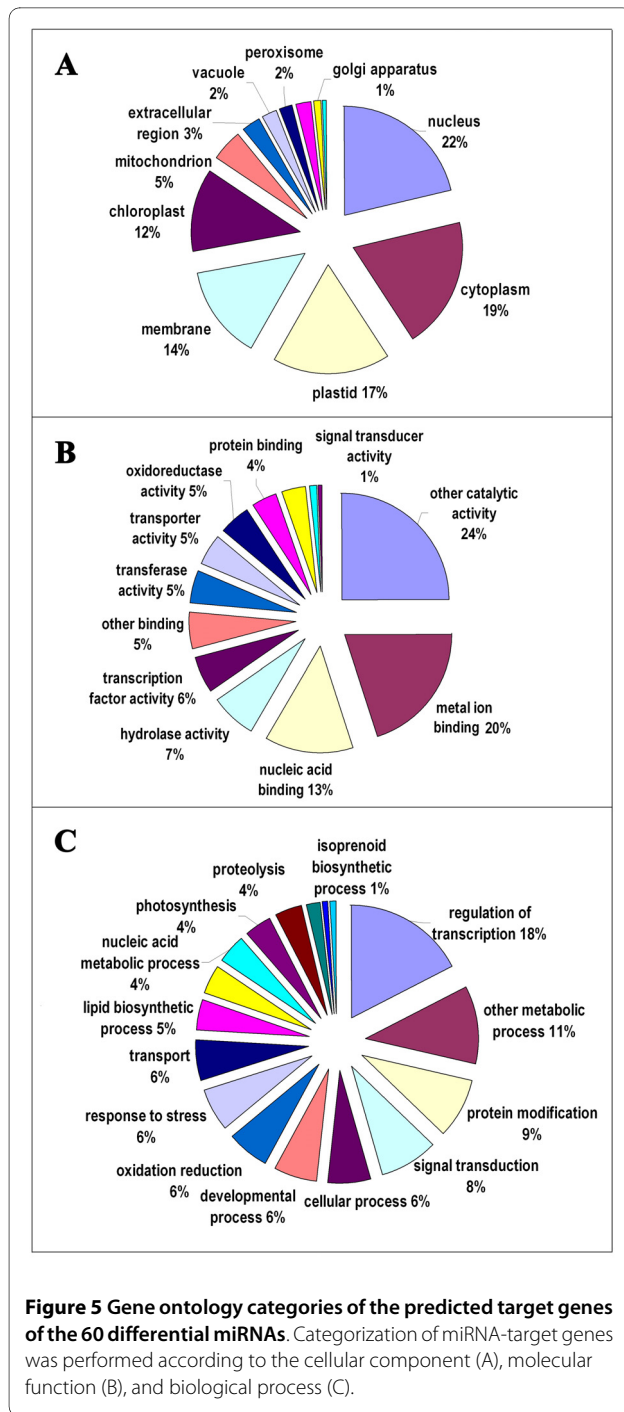
<sup>a</sup> All the 555 miRNA-target genes were used for analysis, detailed gene ID see in Additional file 10.



**Table 4: Top 20 differentially expressed miRNAs between WT and MT**

miRNA	Sequence	WT count <sup>a</sup>	MT count <sup>a</sup>	Fold change	p-value
<b>Higher expression in MT</b>					
csi-miR479	UGUGAUUUGG UUCGGCUCAUC	1341.6	1921.5	1.9	0
csi-miR156a	UGACAGAAGAGA GUGAGCAC	856.2	1656.1	2.6	0
csi-miR167a	UGAAGCUGCCAG CAUGAUCUA	759.6	1967.3	21.6	0
csi-miR171b	CGAGCCGAAUCA AUAUCACUC	28.5	615.4	4.5	0
csi-novel-08	UAGAUAAAGAU GAGAGAAAAA	220.1	1248.5	1.5	0
csi-miR166j	UCUCGGACCAGG CUUCAUCC	1682.9	2078.5	3.1	2.85E-22
<b>Lower expression in MT</b>					
csi-miR482a	UCUCCCUAUGC CUCCAUUCC	103.9	4.2	15.6	0
csi-miR162a	UCGAUAAACCUC UGCAUCCAG	401.0	25.7	12.4	0
csi-miR159a	UUUGGAUUGAA GGGAGCUCUA	307.8	24.8	3.5	0
csi-miR390a	AAGCUCAGGAGG GAUAGCGCC	822.2	235.8	2.5	0
csi-miR164a	UGGAGAAGCAG GGCACGUGCA	1894.0	749.0	1.6	0
csi-miR167d	UGAAGCUGCCAG CAUGAUCUGA	1362.3	825.8	1.4	0
csi-miR172a	AGAAUCUUGAU GAUGCUGCAU	4688.4	3299.7	1.4	0
csi-novel-11	UGGACAGAGAAA UCACGGUCA	591.5	132.7	5.7	0
csi-miR156b	CUGACAGAAGAC AGUGAGCAC	41.3	1.8	2.5	2.86E-27
csi-miR473a	ACUCUCCUCA GGGCUUCGC	168.1	68.1	1.2	2.01E-26
csi-novel-03	UGAAGGGCCUU UCUAGAGCAC	90.5	29.6	8.1	2.87E-20
csi-miR399a	UGCCAAAGGAGA AUUGCCCUG	36.3	4.5	13.8	3.14E-18
csi-miR482c	UCUUGCCCACCC CUCCAUUCC	29.0	2.1	15.6	8.98E-18
Csi-miR164d	UGGAGAAGCAG GGCACGUGCU	49.1	11.1	4.4	7.6E-17

<sup>a</sup> count was normalized as transcripts per million (TPM).



NAs. To the author's knowledge, this is the first report of cloning and characterization of large scale sRNAs and identification of novel miRNAs from sweet orange. A total of 5 and 7 million sRNAs were obtained from WT and MT, respectively; and the size distribution of small RNAs (Figure 2) is quite similar to those published from phylogenetically close plant species such as Arabidopsis [65], with the total percentage of 21-24 nt sRNAs being more than 90% and percentage of 24 nt sRNAs over 50%.

This suggests that the Illumina sequencing data of sRNA libraries was reliable.

### miRNAome in sweet orange

In this study, 85 conserved miRNAs belonging 48 families as well as 12 novel miRNA families were identified in sweet orange. The miRNAs cloned in this study were compared with those from previously reported citrus miRNAs [36,66,67], of them four sweet orange miRNAs were not overlap with this study. These results suggest that at least 101 miRNAs exist in the miRNA transcriptome (miRNAome) of sweet orange. We believe that this is a conservative estimate since only fruit samples were used in this study. The miRNA species in wild type and mutant was different, i.e. 69 miRNAs were detected in WT sample while 66 miRNAs were in MT sample. Moreover, the expression levels of each miRNA varied greatly within the transcriptome, i.e. in MT the abundance ranged from 0.9 TPM for csi-miR1858 to 8437 TPM for csi-miR168. All these data implied that miRNAome, as small RNA transcriptome [38], is distinct and complex though miRNA sequences are believed to be conserved across different plant families [23,29,68].

Take a broader view on the small RNA transcriptome of sweet orange, we noticed that the size distribution pattern of sRNAs from sweet orange (the total percentage of 21-24 nt sRNAs > 90% with 24 nt sRNAs > 50%) is quite similar to Arabidopsis as described earlier [65]; however, the distribution pattern is strikingly different from that in *Pinus contorta*, a conifer species [37] where 24-nt RNAs was less frequently (2.5%) and 21-nt RNAs were more abundant (> 50%). The striking difference also existed when compare the sweet orange small RNA transcriptome with another gymnosperm plant species, Chinese yew (*Taxus chinensis*) [50]. When compared with monocot species rice, the difference is less severe but still exists, i.e. in rice 24-nt sRNAs has the highest percentage (approximately 20%) followed by the 21-nt sRNAs (approximately 10%) [37]. These analyses indicated that the small RNA transcriptome is complicated across plant species and can be significantly different between phylogenetically-distant plant families.

A total of 12 novel miRNAs were identified from sweet orange, as evidenced from the biogenesis characteristics of miRNAs including the stem-loop structure of the pre-miRNAs (Additional file 8) and sequencing of the miRNA\* in the libraries. The expression of all the 12 miRNAs was also detected by qRT-PCR analyses. Additional qRT-PCR analyses of miRNA\*s revealed that some miRNA\* species can accumulated at certain levels in fruits; moreover, the expression pattern of miRNA\*s in WT and MT fruits can be different from that of miRNAs. For example, csi-novel-05 has lower expression in MT (30 TPM) than in WT (90 TPM) as confirmed by qRT-PCR

**Table 5: List of the important KEGG pathways more than 3 miRNA-target genes affiliated**

KEGGa pathway	Genesb	Gene ID	Best E-value
Benzoate degradation	5	<a href="#">EY732960</a> , <a href="#">EY732985</a> , <a href="#">EY748192</a> , <a href="#">TC18796</a> , <a href="#">TC9807</a>	1.00E-125
Carbon fixation	8	<a href="#">EY665527</a> , <a href="#">EY667061</a> , <a href="#">TC11997</a> , <a href="#">TC13950</a> , <a href="#">TC17473</a> , <a href="#">TC24400</a> , <a href="#">EY722705</a> , <a href="#">TC11997</a>	1.00E-128
Ethylbenzene degradation	5	<a href="#">EY732960</a> , <a href="#">EY732985</a> , <a href="#">EY748192</a> , <a href="#">TC18796</a> , <a href="#">TC9807</a>	1.00E-156
Fatty acid metabolism	5	<a href="#">EY732960</a> , <a href="#">EY732985</a> , <a href="#">EY748192</a> , <a href="#">TC18796</a> , <a href="#">TC9807</a>	1.00E-156
Flavonoid biosynthesis	6	<a href="#">EY697662</a> , <a href="#">EY742573</a> , <a href="#">TC14650</a> , <a href="#">TC3846</a> , <a href="#">TC14650</a> , <a href="#">TC3846</a>	6.00E-93
Geraniol degradation	5	<a href="#">EY732960</a> , <a href="#">EY732985</a> , <a href="#">EY748192</a> , <a href="#">TC18796</a> , <a href="#">TC9807</a>	1.00E-156
Glyoxylate and dicarboxylate metabolism	3	<a href="#">EY667061</a> , <a href="#">EY708354</a> , <a href="#">EY722705</a>	2.00E-86
Linolenic acid metabolism	5	<a href="#">EY732960</a> , <a href="#">EY732985</a> , <a href="#">EY748192</a> , <a href="#">TC18796</a> , <a href="#">TC9807</a>	1.00E-122
Methane metabolism	3	<a href="#">EY697129</a> , <a href="#">EY724821</a> , <a href="#">TC20228</a>	1.00E-68
Pentose phosphate pathway	3	<a href="#">TC13950</a> , <a href="#">TC17473</a> , <a href="#">TC15377</a>	5.00E-58
Phenylalanine metabolism	4	<a href="#">EY664657</a> , <a href="#">EY697129</a> , <a href="#">NP13097244</a> , <a href="#">TC20228</a>	1.00E-167
Photosynthesis - antenna proteins	5	<a href="#">EY672715</a> , <a href="#">EY682256</a> , <a href="#">EY701112</a> , <a href="#">EY701374</a> , <a href="#">EY756276</a>	1.00E-102
Ribosome	3	<a href="#">EY681302</a> , <a href="#">EY752996</a> , <a href="#">TC18634</a>	2.00E-91
Terpenoid biosynthesis	3	<a href="#">EY730974</a> , <a href="#">TC363</a> , <a href="#">EY752486</a>	6.00E-84
Ubiquitin mediated proteolysis	3	<a href="#">EY656660</a> , <a href="#">EY665372</a> , <a href="#">EY752038</a>	2.00E-83
Valine leucine and isoleucine degradation	6	<a href="#">EY732985</a> , <a href="#">EY732960</a> , <a href="#">EY748192</a> , <a href="#">TC18796</a> , <a href="#">TC9807</a> , <a href="#">TC5927</a>	1.00E-156

<sup>a</sup>KEGG = Kyoto Encyclopedia of Genes and Genomes.

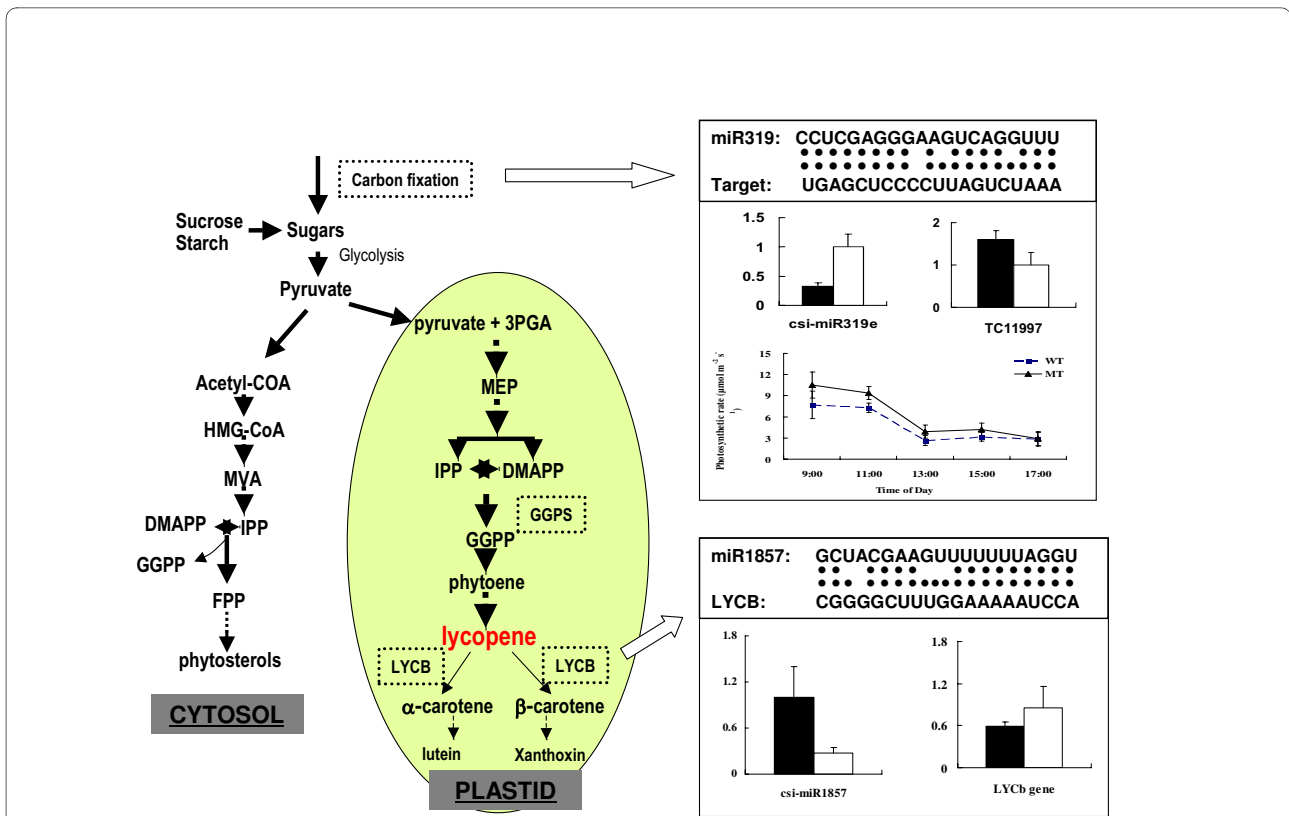
<sup>b</sup>The putative target genes of miRNA genes which differentially expressed significant at 0.05 level between the mutant and wild type

analysis; in contrast, *csi-novel-05\** has higher expression in MT (50 TPM) than in WT (26 TPM) which was also confirmed by qRT-PCR. The differential accumulation of miRNAs and miRNA\*s may be due to the different activities of key enzyme in the sRNA-pathway, as shown in rice that the differential expression of AGO genes in different tissues might cause the change of miRNA\*s abundance [44].

#### miRNA-targets are hot spots for generating sRNAs

The relationship of sRNA genes and the miRNA target genes is one of the hot topics in the phenomenon of

"miRNA-associated transitivity" [69-71]. In this study, 13.7% (WT) and 17.5% (MT) of the miRNA putative target genes generated sRNAs. This data implied that miRNA target gene could be hot spots for generating sRNAs. Moreover, high percentage of antisense sRNAs was associated with the miRNA-target genes, with 11.7% and 12.6% for WT and MT, respectively. The antisense sRNAs/transcripts can function in post-transcriptional gene silencing [62,72]. The post-transcriptional gene silencing is mainly caused by formation of dsRNA duplexes; this could be supported from the observation that 75% of the antisense sRNAs have their cognate sense



**Figure 6** The model for the biological processes possibly regulated by miRNAs in the red-flesh mutant of sweet orange. The pathway was modified based on Xu et al. (2009) [13]. The two biological processes, the carotenogenesis and photosynthesis, were high ranked processes that differential miRNA target genes involved in by gene ontology analysis. The data of photosynthetic rate were extracted from our previous in situ photosynthesis analysis [13]. The expression profiles of miRNA319 and miR1857 are complementary to the profiles of the corresponding miRNA-target genes (phosphoenolpyruvate carboxylase [TC11997]) and lycopene  $\beta$ -cyclase [LYCB]), implied the possible roles of these miRNAs in regulation the biological processes on posttranscriptional way.

transcripts in WT and 64% happens in MT sample in this study.

### miRNAs in plant fruits

Up to now, miRNA identification in fruits using a cloning approach is rare except in tomato [23] and grape [60]. In this study, cloning from sweet orange fruit led to the identification of 48 conserved miRNA families. Analysis of them revealed that 6 miRNA families are expressed in fruits of the three plant species; but they are not fruit-specific. Furthermore, the comparison between the two fruit crops (85 miRNAs from sweet orange and 140 miRNAs from grapevine) showed that 8 miRNA families are expressed in fruits of both crops, and only miR535 is specific to the two species. These analyses suggested that expression and function of fruit-specific miRNAs across plant species is lacking, and miRNAs involved in fruit ripening may overlap with that function in other organs, or

the same miRNAs in different organs may have different functions.

### miRNAs possibly involved in the regulation of lycopene accumulation in the sweet orange red-flesh mutant

The characterization and comparative profiling of entire sets of small RNAs (Small RNA transcriptome), especially the microRNA transcriptome, lays the foundation to unravel the complex miRNA-mediated regulatory networks controlling lycopene accumulation in the red-flesh sweet orange mutant. In this study, a number of miRNAs have been identified to be differentially expressed between 'Anliu' sweet orange (WT) and its red-flesh mutant (MT). Compared with WT, 16 known miRNA genes and one novel miRNA genes were expressed exclusively in MT. On the other hand, we found 19 known miRNAs and two novel miRNAs were WT-specific. Moreover, a total of 51 known miRNAs and 9 novel miR-

NAs were found to exhibit significant expression changes between MT and WT.

Target prediction of these differential miRNAs could provide information on the biological processes regulated by miRNA. Putative targets were predicted by psRNA-Target (the updated version of miRU) [56] and web miRNA designer (WMD3) [57]. This strategy was also used in Arabidopsis for the identification of nutrient-responsive miRNAs and their potential targets [24], though further direct validation is necessary to elucidate the true in-vivo targets. However, the annotations of these potential miRNA target genes did provide an alternate view of gene regulation of the red-flesh trait formation in MT (Additional file 11). We found that three groups of predicted miRNA-target genes are most possibly involved in the mutant trait formation of lycopene accumulation. The first interesting observation is that two potential targets (EY752486 by *csi-miR167* and TC5 by *csi-miR1857*) are the key enzymes of the carotenoid biosynthesis pathway (Additional file 11). They encode geranylgeranyl pyrophosphate synthase (GGPS), a key enzyme for the production of prerequisite molecules for carotenoids accumulation, and lycopene  $\beta$ -cyclase (LYCb), a rate-limiting enzyme in the conversion of lycopene to downstream cyclic carotenes. The expression of LYCb and *csi-miR1857* also indicated possible interaction between them because of the complementary expression profile (Figure 6). And we also detected the GGPS gene expression, but failed due to the incorrect sequence information in the current public database. The second interesting target genes are APETALA2 (AP2) transcription factors, including EY726563 by *csi-novel-03*, and TC7810 by *csi-miR172a*. In Arabidopsis, one AP2 protein (designated as RAP2.2) has been experimentally validated to interact with phytoene synthase (PSY) gene, which encodes a rate-limiting enzyme for the synthesis of carotenoids. The RAP2.2 gene was also believed to be regulated by miRNA because of the large discrepancy between mRNA and protein levels in the overexpression lines [15]. The data from this study also supported the proposition, and indicated that RAP2.2 (encoded by TC7810 gene in this study) is a potential target of miRNA172a in sweet orange. The third group of target genes is those involved in photosynthesis or carbon fixation (Table 5). The potential target genes associated with photosynthesis fits well with the GO results with large number of genes that are located in chloroplast or plastid. Moreover, our photosynthesis-related gene expression analysis and in situ detection of the photosynthetic rate confirmed that photosynthesis in MT is stronger than that in WT (Figure 6). The result is quite consistent with our gas chromatographic analyses that the sugars in juice sacs are higher in MT than in WT [11]. Tomato *hp-2* mutant which manifested with increased carotenoids,

primarily lycopene, showed that genes related with photosynthesis are consistently up-regulated throughout fruit ripening [73]. All these evidences implied the important role of photosynthesis in the accumulation of lycopene. Taken together, the annotation of the potential targets of differentially expressed miRNA genes between WT and MT indicated that several miRNA-directed biological processes including carotenoid biosynthesis, transcription regulation, and photosynthesis may be important for the regulation of lycopene accumulation in the sweet orange red-flesh mutant.

## Conclusions

Our study provides a large scale cloning and characterization of miRNAs in sweet orange. The interpretation of the small RNA data uncovered a total of 85 known and 12 novel miRNAs. Meanwhile, comparative profiling of these miRNAs between a spontaneous red-flesh mutant and its wild type revealed that 51 known and 9 novel miRNAs are differentially expressed. Target prediction of the 60 differential miRNAs indicated that two key genes of the carotenoid biosynthesis, geranylgeranyl pyrophosphate synthase (GGPS) and lycopene  $\beta$ -cyclase (LYCb), are under miRNA regulation. Functional annotation of the 418 target genes of differential miRNA revealed that high ranked biological processes are transcription regulation, protein modification and photosynthesis. The expression of miRNAs and target genes involved in photosynthesis and carotenogenesis were further confirmed. Taken together, this study provides new insight into the miRNAome and possible miRNA-directed molecular processes that regulate lycopene accumulation in the red-flesh sweet orange mutant.

## Methods

### Plant materials and RNA preparation

The red-flesh mutant 'Hong Anliu' and the wild type 'Anliu' sweet orange (*Citrus sinensis* L. Osbeck), cultivated at the Institute of Citrus Research located in Guilin, Guangxi Province, China, were used as materials. The mutant 'Hong Anliu' sweet orange (MT) is a spontaneous bud mutation of wild type 'Anliu' sweet orange (WT); and they are with isogenic background as revealed from our previous physiological traits and molecular marker evaluations [11]. Same sampling strategy as previously published [13] was employed. Fruit samples were harvested at 170 days after flowering (DAF) that is the critical stage for the flesh color transition. Three different trees were used, and 10 representative fruits from each tree were collected. All samples were separated into peel and pulp, immediately frozen in liquid nitrogen and kept at  $-80^{\circ}\text{C}$  until use. Total RNA was extracted according to Liu et al. [74].

### Small RNA library construction and sequencing

Small RNA library construction and sequencing was performed as described by Yao et al. [49]. Briefly, small RNAs with 16-30 nt in length were isolated from the total RNA of WT and MT samples by size fractionation with 15% Tris-borate-EDTA urea polyacrylamide gel (TBU). After extraction from the gel, sRNA was dephosphorylated by alkaline phosphatase (New England Biolabs Inc.). Then the small RNAs were ligated sequentially to 5' and 3' RNA/DNA chimeric oligonucleotide adaptors (Illumina). The ligation products were purified from 10% TBE urea polyacrylamide gel. Reverse transcription was performed with the ligated sRNAs as templates. The obtained cDNAs were sequenced using the Illumina 1 G Genome Analyzer by Beijing Genomics Institute (BGI) (Shenzhen, China).

### Small RNA annotation and microRNA identification

Small RNA sequences were extracted from raw reads matching both the last 7 nt of the 5'-adaptor and the first 7 nt of the 3'-adaptor as described by Moxon et al. [23]. After trimming off the adaptors, sequences between 18 and 30 bp in length and reads more than 3 were used to search the Rfam database <http://www.sanger.ac.uk/Software/Rfam> by BLASTN analysis to discard rRNA, tRNA, snRNA, and snoRNA. The remaining sequences were compared to known plant miRNAs from miRBase v13.0 [75] using BLASTN algorithm to search conserved miRNAs. Sequences in our libraries with identical or with four or fewer mismatches to currently known miRNAs from other plant species were regarded as potential conserved miRNAs. The potential miRNA sequences were used for BLASTN search against sweet orange unigene dataset and 1.2× coverage sequence of sweet orange (from NCBI), the perfectly matched sequences were used for fold-back structure prediction by Mfold program [59].

To identify non-conserved miRNAs from sweet orange, the remained sRNAs were used for BLASTN search against sweet orange ESTs from TIGR gene index database [http://compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/GO\\_browser.pl?species=Orange&gi\\_dir=csgi](http://compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/GO_browser.pl?species=Orange&gi_dir=csgi) and 1.2× coverage sequence of sweet orange (from NCBI), since the sweet orange genome is not fully sequenced. The sequences of perfect match with sRNA sequences were used for fold-back structure prediction by Mfold program [59]. The sRNAs that have stem-loop precursors were regarded as putative non-conserved miRNAs. Furthermore, we searched miRNA\* sequences (complementary to miRNA in the precursor molecule) in the sRNA libraries. Only those with miRNA-miRNA\* duplexes were regarded as novel miRNAs.

Significance level of the difference of miRNA frequency and abundance between WT and MT was analyzed using Z-score method according to Kal et al. [76]. The signifi-

cance of differentially expressed miRNAs were set at the  $p$  value < 0.01.

### miRNA expression confirmation by stem-loop qRT-PCR

The detection of miRNA expression was performed using stem-loop qRT-PCR method according to Chen et al. [53] and Varkonyi-Gasic et al. [54]. Stem-loop primers for reverse transcription and primers for quantitative PCR were listed in Additional file 12. Total RNA was reverse-transcribed using SuperScript III Reverse Transcriptase (Invitrogen, US) by a pulse reverse transcription program [54]. Real-time quantitative PCR (qPCR) was performed on the ABI 7500 Real Time System (PE Applied Biosystems, Foster City, CA, USA) using *actin* gene as endogenous control according to Liu et al. [11]. Briefly, the primers for miRNAs and *actin* were diluted in the SYBER GREEN PCR Master Mix (PE Applied Biosystems) and 20  $\mu$ l of the reaction mix was added to each well. Reactions were performed by an initial incubation at 50°C for 2 min and at 95°C for 1 min, and then cycled at 95°C for 15 s and 60°C for 1 min for 40 cycles. Output data was generated by the instrument on-board software Sequence Detector Version 1.3.1 (PE Applied Biosystems). The electrophoresis of the PCR products was performed on 3% agarose gel for 1 hr under 70 V voltage. The marker contained two fragments, 50 and 60 bp in length, the sequences information was: GTCGTATCCAGTGCAGGGTCCGAGG-TATTCGCACTGGATACGACCCGATG and CGGCG AAGATTGAAGTGAAGGTGCAGGGTCCGAGGTATTCGCACTGGATACGACTTCCGA respectively.

### Target prediction of miRNAs

Target prediction of miRNAs from sweet orange was performed based on methods described by Pant et al. [24]. Two publicly available algorithms, miRU [56] and WMD3 [57] were used in the prediction. The sweet orange EST sequences from TIGR database as described earlier were used to predict the targets. The hybridization energy of each miRNA-predicted target duplex ( $\Delta G$ ) was determined from WMD3 output.

### Functional assignments of the potential targets of differentially expressed miRNAs

To investigate biological processes possibly regulated by miRNAs, assignments of putative functions to potential target genes of differentially expressed miRNAs were performed using annot8r program, which was run locally to BLAST against a reference database that stores UniProt entries, their associated Gene Ontology (GO), Enzyme Commission (EC) and Kyoto Encyclopaedia of Genes and Genomes (KEGG) annotations [77]. The GO categorization results were expressed as three independent hierarchies for biological process, cellular component, and molecular function [78]. The biological interpretation of

the potential targets was further completed by assigning to metabolic pathways using KEGG [79].

#### The expression detection of two microRNAs target genes

The expression levels of one carotenoid biosynthesis gene (lycopene  $\beta$ -cyclase, LYCb) and one carbon fixation-related gene (phosphoenolpyruvate carboxylase, TC11997) were determined on the ABI 7500 Real Time System (PE Applied Biosystems, Foster City, CA, USA) using *actin* gene as endogenous control according to Liu et al. [11]. The primers are as following: TC11997: forward, GAGATGGTGCAGGCCAAGTT and reverse GTTGAGGAGGTCGCATGGTT; LYCb: forward, GGC-TATATGGTGGCAAGGACTT and reverse, CAGAATT-GAGGCTTCGAACGA. The data of in situ rates of photosynthesis in the mutant and the wildtype were extracted from our previous data [13].

#### Data access

The small RNA sequence data from this study have been submitted to Gene Expression Omnibus (GEO) under accession No. GSE18207 at website: <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=bvncfemmqckehi&acc=GSE18207>

#### Additional material

**Additional file 1 WT-derived sRNAs from coding RNA.** This file contains the information of WT-derived sRNAs which were mapped on protein-coding EST genes from sweet orange (*Citrus sinensis*).

**Additional file 2 MT-derived sRNAs from coding RNA.** This file contains the information of MT-derived sRNAs which were mapped on protein-coding EST genes from sweet orange (*Citrus sinensis*).

**Additional file 3 WT-derived sRNAs from non-coding RNA.** The file lists WT-derived sRNAs which were mapped on non-protein coding RNAs (tRNA, snRNA, snoRNA and rRNA).

**Additional file 4 MT-derived sRNAs from non-coding RNA.** The file lists MT-derived sRNAs which were mapped on non-protein coding RNAs (tRNA, snRNA, snoRNA and rRNA).

**Additional file 5 Known citrus miRNAs.** Known citrus miRNAs and their expression in WT and MT sweet oranges.

**Additional file 6 Fold-back structures for known miRNA from sweet orange (*Citrus sinensis*).** Precursor sequences for known miRNAs from sweet orange were shown in black letters with miRNA and miRNA\* (The sequence complementary to miRNA in the fold-back structure) sequences highlighted in yellow and pink, respectively. Precursor secondary structures and dG value were produced using the mfold software <http://mfold.bio-info.rpi.edu/>.

**Additional file 7 Target prediction of citrus known miRNAs.** This file contains the predicted target genes for known miRNA genes in sweet orange.

**Additional file 8 Fold-back structures for novel miRNAs from sweet orange (*Citrus sinensis*).** Precursor sequences for novel miRNA from sweet orange were shown in black letters with miRNA and miRNA\* (The sequence complementary to miRNA in the fold-back structure) sequences highlighted in yellow and pink, respectively. Precursor secondary structures and dG value were produced using the mfold software <http://mfold.bio-info.rpi.edu/>.

**Additional file 9 Target prediction of citrus novel miRNAs.** This file contains the predicted target genes for novel miRNA genes in sweet orange.

**Additional file 10 miRNA target genes are hotspots for generation small RNAs.** This file contains the information of citrus sRNA loci located in miRNA-target genes in both sense and antisense way.

**Additional file 11 The computational annotation of target genes of differential miRNA genes.** This file shows the annotation information of citrus miRNA-target genes; the annotation was performed using arbidopsis homologous gene and Uniprot database.

**Additional file 12 The primer sequence information.** This file lists the primers used for miRNA expression detection by stem-loop qRT-PCR.

#### Abbreviations

AGO: ARGONAUTE; AP2: APETAL2; bZIP: Basic zipper; DAF: Days after flowering; GO: Gene ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; miRNA: microRNA; NAC: NAM, ATAF, and CUC; PEPC: phosphoenolpyruvate carboxylase; rRNA: Ribosomal RNA; Rubisco: Ribulose-1,5-bisphosphate carboxylase oxygenase; SBP: Squamosa promoter-binding protein; snRNA: Small nuclear RNA; snoRNA: Small nucleolar RNA; sRNA: small RNA; siRNA: short interfering RNA; TPM: Transcripts per million; tRNA: transfer RNA.

#### Authors' contributions

QX, YLL, ADZ and KQY are responsible for generating the sRNA data and for interpretation of the results. YLL and JLY carried out qRT-PCR experiments. QX drafted the manuscript. XMW and WWG participated in research design and statistical analyses. XXD proposed and supervised the research. All authors read and approved the final manuscript.

#### Acknowledgements

We thank Dr. Hanhui Kuang from Huazhong Agricultural University (China) for their helpful suggestions on this study. This work was supported by the Natural Science Foundation of China (NSFC Nos. 30921002, 30830078 and 30800745), Specialized Research Fund for the Doctoral Program of Ministry of Education of China (No.200805041024), and Hubei Provincial Natural Science Foundation (2008CDB083).

#### Author Details

<sup>1</sup>Key Laboratory of Horticultural Plant Biology of Ministry of Education, Huazhong Agricultural University, Wuhan 430070, China and <sup>2</sup>National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan 430070, China

Received: 10 November 2009 Accepted: 17 April 2010

Published: 17 April 2010

#### References

1. Spiegel-Roy P, Goldschmidt EE: **Biology of Citrus.** Cambridge, UK: Cambridge University Press; 1996.
2. Jayaprakasha GK, Patil BS: **In vitro evaluation of the antioxidant activities in fruit extracts from citron and blood orange.** *Food Chem* 2007, **101**:410-418.
3. Bartley GE, Scolnik PA: **Plant carotenoids: pigments for photoprotection, visual attraction, and human health.** *Plant Cell* 1995, **7**:1027-1038.
4. Rouseff R, Raley L, Hofsommer HJ: **Application of diode array detection with a C-30 reversed phase column for the separation and identification of saponified orange juice carotenoids.** *J Agric Food Chem* 1996, **44**:2176-2181.
5. Lee HS, Castle WS: **Seasonal changes of carotenoid pigments and color in Hamlin, Earlygold, and budd blood orange juices.** *J Agric Food Chem* 2001, **49**:877-882.
6. Xu CJ, Fraser PD, Wang WJ, Bramley PM: **Differences in the carotenoid content of ordinary citrus and lycopene-accumulating mutants.** *J Agric Food Chem* 2006, **54**:5474-5481.
7. Kato M, Ikoma Y, Matsumoto H, Sugiura M, Hyodo H, Yano M: **Accumulation of carotenoids and expression of carotenoid biosynthetic genes during maturation in citrus fruit.** *Plant Physiol* 2004, **134**:824-837.
8. Schwarz S, Obermuller-Jevic UC, Hellmis E, Koch W, Jacobi G, Biesalski HK: **Lycopene inhibits disease progression in patients with benign prostate hyperplasia.** *J Nutr* 2008, **138**:49-53.
9. Monselise SP, Halevy AH: **Detection of lycopene in pink orange fruit.** *Science* 1961, **133**:1478.

10. Saunt J: **Citrus Varieties of the World**. Norwich, England: Sinclair International Ltd; 2000.
11. Liu Q, Xu J, Liu YZ, Zhao XL, Deng XX, Guo LL, Gu JQ: **A novel bud mutation that confers abnormal patterns of lycopene accumulation in sweet orange fruit (*Citrus sinensis* L. Osbeck)**. *J Exp Bot* 2007, **58**:4161-4171.
12. Liu Q, Zhu A, Chai LJ, Zhou WJ, Yu KQ, Ding J, Xu J, Deng XX: **Transcriptome analysis of a spontaneous mutant in sweet orange [*Citrus sinensis* (L.) Osbeck] during fruit development**. *J Exp Bot* 2009, **60**:801-813.
13. Xu Q, Yu KQ, Zhu AD, Ye JL, Liu Q, Zhang JC, Deng X: **Comparative transcripts profiling reveals new insight into molecular processes regulating lycopene accumulation in a sweet orange (*Citrus sinensis*) red-flesh mutant**. *BMC Genomics* 2009, **10**:540.
14. Pan ZY, Liu Q, Yun Z, Guan R, Zeng WF, Xu Q, Deng XX: **Comparative proteomics of a lycopene-accumulating mutant reveals the important role of oxidative stress on carotenogenesis in sweet orange (*Citrus sinensis* [L.] Osbeck)**. *Proteomics* 2009, **9**:5455-5470.
15. Welsch R, Maass D, Voegel T, Dellapenna D, Beyer P: **Transcription factor RAP2.2 and its interacting partner SINAT2: stable elements in the carotenogenesis of Arabidopsis leaves**. *Plant Physiol* 2007, **145**:1073-1085.
16. Bartel DP: **MicroRNAs: target recognition and regulatory functions**. *Cell* 2009, **136**:215-233.
17. Vaucheret H: **Post-transcriptional small RNA pathways in plants: mechanisms and regulations**. *Genes Dev* 2006, **20**:759-771.
18. Jones-Rhoades MW, Bartel DP: **Computational identification of plant MicroRNAs and their targets, including a stress-induced miRNA**. *Mol Cell* 2004, **14**:787-799.
19. Sunkar R, Zhu JK: **Novel and stress-regulated microRNAs and other small RNAs from Arabidopsis**. *Plant Cell* 2004, **16**:2001-2019.
20. Lu S, Sun YH, Shi R, Clark C, Li L, Chiang VL: **Novel and mechanical stress-responsive MicroRNAs in *Populus trichocarpa* that are absent from Arabidopsis**. *Plant Cell* 2005, **17**:2186-2203.
21. Mallory AC, Vaucheret H: **Functions of microRNAs and related small RNAs in plants**. *Nat Genet* 2006, **38**:S31-36.
22. Aung K, Lin SI, Wu CC, Huang YT, Su CL, Chiou TJ: **pho2, a phosphate overaccumulator, is caused by a nonsense mutation in a MicroRNA399 target gene**. *Plant Physiol* 2006, **141**:1000-1011.
23. Moxon S, Jing R, Szittya G, Schwach F, Rusholme Pilcher RL, Moulton V, Dalmay T: **Deep sequencing of tomato short RNAs identifies microRNAs targeting genes involved in fruit ripening**. *Genome Res* 2008, **18**:1602-1609.
24. Pant BD, Musialak-Lange M, Nuc P, May P, Buhtz A, Kehr J, Walther D, Scheible WR: **Identification of nutrient-responsive Arabidopsis and rapeseed microRNAs by comprehensive real-time polymerase chain reaction profiling and small RNA sequencing**. *Plant Physiol* 2009, **150**:1541-1555.
25. Ramachandran V, Chen X: **Small RNA metabolism in Arabidopsis**. *Trends Plant Sci* 2008, **13**:368-374.
26. Llave C, Kasschau KD, Rector MA, Carrington JC: **Endogenous and silencing-associated small RNAs in plants**. *Plant Cell* 2002, **14**:1605-1619.
27. Lindow M, Krogh A: **Computational evidence for hundreds of non-conserved plant microRNAs**. *BMC Genomics* 2005, **6**:119.
28. Rajagopalan R, Vaucheret H, Trejo J, Bartel DP: **A diverse and evolutionarily fluid set of microRNAs in *Arabidopsis thaliana***. *Genes Dev* 2006, **20**:3407-3425.
29. Jones-Rhoades MW, Bartel DP, Bartel B: **MicroRNAs and their regulatory roles in plants**. *Annu Rev Plant Biol* 2006, **57**:19-53.
30. Wang JF, Zhou H, Chen YQ, Luo QJ, Qu LH: **Identification of 20 microRNAs from *Oryza sativa***. *Nucleic Acids Res* 2004, **32**:1688-1695.
31. Sunkar R, Girke T, Jain PK, Zhu JK: **Cloning and characterization of microRNAs from rice**. *Plant Cell* 2005, **17**:1397-1411.
32. Bonnet E, Wuyts J, Rouze P, Peer Y Van de: **Detection of 91 potential conserved plant microRNAs in *Arabidopsis thaliana* and *Oryza sativa* identifies important target genes**. *Proc Natl Acad Sci USA* 2004, **101**:11511-11516.
33. Adai A, Johnson C, Mlotshwa S, Archer-Evans S, Manocha V, Vance V, Sundaresan V: **Computational prediction of miRNAs in *Arabidopsis thaliana***. *Genome Res* 2005, **15**:78-91.
34. Zhang BH, Pan XP, Wang QL, Cobb GP, Anderson TA: **Identification and characterization of new plant microRNAs using EST analysis**. *Cell Res* 2005, **15**:336-360.
35. Gleave AP, Ampomah-Dwamena C, Berthold S, Dejnopratt S, Karunairetnam S, Nain B, Wang Y-Y, Crowhurst RN, MacDiarmid RM: **Identification and characterization of primary microRNAs from apple (*Malus domestica* cv. Royal Gala) expressed sequence tags**. *Tree Genet Genomes* 2008, **4**:343-358.
36. Song C, Fang J, Li X, Liu H, Thomas Chao C: **Identification and characterization of 27 conserved microRNAs in citrus**. *Planta* 2009, **230**:671-685.
37. Morin RD, Aksay G, Dolgosheina E, Ebhardt HA, Magrini V, Mardis ER, Sahinalp SC, Unrau PJ: **Comparative analysis of the small RNA transcriptomes of *Pinus contorta* and *Oryza sativa***. *Genome Res* 2008, **18**:571-584.
38. Lu C, Tej SS, Luo S, Haudenschild CD, Meyers BC, Green PJ: **Elucidation of the small RNA component of the transcriptome**. *Science* 2005, **309**:1567-1569.
39. Lu C, Kulkarni K, Souret FF, MuthuVallippan R, Tej SS, Poethig RS, Henderson IR, Jacobsen SE, Wang W, Green PJ, et al.: **MicroRNAs and other small RNAs enriched in the Arabidopsis RNA-dependent RNA polymerase-2 mutant**. *Genome Res* 2006, **16**:1276-1288.
40. Fahlgren N, Howell MD, Kasschau KD, Chapman EJ, Sullivan CM, Cumbie JS, Givan SA, Law TF, Grant SR, Dangl JL, et al.: **High-throughput sequencing of Arabidopsis microRNAs: evidence for frequent birth and death of MIRNA genes**. *PLoS One* 2007, **2**:e219.
41. Nobuta K, Venu RC, Lu C, Belo A, Vemaraju K, Kulkarni K, Wang W, Pillay M, Green PJ, Wang GL, et al.: **An expression atlas of rice mRNAs and small RNAs**. *Nat Biotechnol* 2007, **25**:473-477.
42. 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.
43. Sunkar R, Zhou X, Zheng Y, Zhang W, Zhu JK: **Identification of novel and candidate miRNAs in rice by high throughput sequencing**. *BMC Plant Biol* 2008, **8**:25.
44. Xue LJ, Zhang JJ, Xue HW: **Characterization and expression profiles of miRNAs in rice seeds**. *Nucleic Acids Res* 2009, **37**:916-930.
45. Axtell MJ, Snyder JA, Bartel DP: **Common functions for diverse small RNAs of land plants**. *Plant Cell* 2007, **19**:1750-1769.
46. Barakat A, Wall K, Leebens-Mack J, Wang YJ, Carlson JE, dePamphilis CW: **Large-scale identification of microRNAs from a basal eudicot (*Eschscholzia californica*) and conservation in flowering plants**. *Plant J* 2007, **51**:991-1003.
47. Barakat A, Wall PK, Diloreto S, Depamphilis CW, Carlson JE: **Conservation and divergence of microRNAs in *Populus***. *BMC Genomics* 2007, **8**:481.
48. Molnar A, Schwach F, Studholme DJ, Thuenemann EC, Baulcombe DC: **miRNAs control gene expression in the single-cell alga *Chlamydomonas reinhardtii***. *Nature* 2007, **447**:1126-1129.
49. Yao Y, Guo G, Ni Z, Sunkar R, Du J, Zhu JK, Sun Q: **Cloning and characterization of microRNAs from wheat (*Triticum aestivum* L.)**. *Genome Biol* 2007, **8**:R96.
50. Qiu D, Pan X, Wilson IW, Li F, Liu M, Teng W, Zhang B: **High throughput sequencing technology reveals that the taxoid elicitor methyl jasmonate regulates microRNA expression in Chinese yew (*Taxus chinensis*)**. *Gene* 2009, **436**:37-44.
51. 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.
52. Ambros V, Bartel B, Bartel DP, Burge CB, Carrington JC, Chen XM, Dreyfuss G, Eddy SR, Griffiths-Jones S, Marshall M, et al.: **A uniform system for microRNA annotation**. *RNA* 2003, **9**:277-279.
53. Chen C, Ridzon DA, Broomer AJ, Zhou Z, Lee DH, Nguyen JT, Barbisin M, Xu NL, Mahavakar VR, Andersen MR, et al.: **Real-time quantification of microRNAs by stem-loop RT-PCR**. *Nucleic Acids Res* 2005, **33**:e179.
54. Varkonyi-Gasic E, Wu R, Wood M, Walton EF, Hellens RP: **Protocol: a highly sensitive RT-PCR method for detection and quantification of microRNAs**. *Plant Methods* 2007, **3**:12.
55. Feng J, Wang K, Liu X, Chen S, Chen J: **The quantification of tomato microRNAs response to viral infection by stem-loop real-time RT-PCR**. *Gene* 2009, **437**:14-21.



56. Zhang Y: **miRU: an automated plant miRNA target prediction server.** *Nucleic Acids Res* 2005, **33**:W701-704.
57. Ossowski S, Schwab R, Weigel D: **Gene silencing in plants using artificial microRNAs and other small RNAs.** *Plant J* 2008, **53**:674-690.
58. Moxon S, Schwach F, Dalmay T, Maclean D, Studholme DJ, Moulton V: **A toolkit for analysing large-scale plant small RNA datasets.** *Bioinformatics* 2008, **24**:2252-2253.
59. Zuker M: **Mfold web server for nucleic acid folding and hybridization prediction.** *Nucleic Acids Res* 2003, **31**:3406-3415.
60. Carra A, Mica E, Gambino G, Pindo M, Moser C, Pe ME, Schubert A: **Cloning and characterization of small non-coding RNAs from grape.** *Plant J* 2009, **59**:750-763.
61. Arazi T, Talmor-Neiman M, Stav R, Riese M, Huijser P, Baulcombe DC: **Cloning and characterization of micro-RNAs from moss.** *Plant J* 2005, **43**:837-848.
62. Luo QJ, Samanta MP, Koksall F, Janda J, Galbraith DW, Richardson CR, Ouyang F, Rock CD: **Evidence for antisense transcription associated with microRNA target mRNAs in Arabidopsis.** *PLoS Genet* 2009, **5**:e1000457.
63. Morin RD, O'Connor MD, Griffith M, Kuchenbauer F, Delaney A, Prabhu AL, Zhao Y, McDonald H, Zeng T, Hirst M, *et al.*: **Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells.** *Genome Res* 2008, **18**:610-621.
64. Hobert O: **Gene regulation by transcription factors and microRNAs.** *Science* 2008, **319**:1785-1786.
65. Henderson IR, Zhang X, Lu C, Johnson L, Meyers BC, Green PJ, Jacobsen SE: **Dissecting Arabidopsis thaliana DICER function in small RNA processing, gene silencing and DNA methylation patterning.** *Nat Genet* 2006, **38**:721-725.
66. Zhang B, Pan X, Cannon CH, Cobb GP, Anderson TA: **Conservation and divergence of plant microRNA genes.** *Plant J* 2006, **46**:243-259.
67. Sunkar R, Jagadeeswaran G: **In silico identification of conserved microRNAs in large number of diverse plant species.** *BMC Plant Biol* 2008, **8**:37.
68. Voinnet O: **Origin, biogenesis, and activity of plant microRNAs.** *Cell* 2009, **136**(4):669-687.
69. Axtell MJ, Jan C, Rajagopalan R, Bartel DP: **A two-hit trigger for siRNA biogenesis in plants.** *Cell* 2006, **127**:565-577.
70. Ronemus M, Vaughn MW, Martienssen RA: **MicroRNA-targeted and small interfering RNA-mediated mRNA degradation is regulated by argonaute, dicer, and RNA-dependent RNA polymerase in Arabidopsis.** *Plant Cell* 2006, **18**:1559-1574.
71. Allen E, Xie ZX, Gustafson AM, Carrington JC: **microRNA-directed phasing during trans-acting siRNA biogenesis in plants.** *Cell* 2005, **121**:207-221.
72. He Y, Vogelstein B, Velculescu VE, Papadopoulos N, Kinzler KW: **The antisense transcriptomes of human cells.** *Science* 2008, **322**:1855-1857.
73. Kolotilin I, Koltai H, Tadmor Y, Bar-Or C, Reuveni M, Meir A, Nahon S, Shlomo H, Chen L, Levin I: **Transcriptional profiling of high pigment-2dg tomato mutant links early fruit plastid biogenesis with its overproduction of phytonutrients.** *Plant Physiol* 2007, **145**:389-401.
74. Liu Y, Liu Q, Tao N, Deng X: **Efficient isolation of RNA from fruit peel and pulp of ripening navel orange (Citrus sinensis Osbeck).** *J Huazhong Agric Univ* 2006, **25**:300-304.
75. Griffiths-Jones S: **miRBase: the microRNA sequence database.** *Methods Mol Biol* 2006, **342**:129-138.
76. Kal AJ, van Zonneveld AJ, Benes V, Berg M van den, Koerkamp MG, Albermann K, Strack N, Ruijter JM, Richter A, Dujon B, *et al.*: **Dynamics of gene expression revealed by comparison of serial analysis of gene expression transcript profiles from yeast grown on two different carbon sources.** *Mol Biol Cell* 1999, **10**:1859-1872.
77. Schmid R, Blaxter ML: **annot8r: GO, EC and KEGG annotation of EST datasets.** *BMC Bioinformatics* 2008, **9**:180.
78. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, *et al.*: **Gene Ontology: tool for the unification of biology.** *Nat Genet* 2000, **25**:25-29.
79. Kanehisa M, Goto S: **KEGG: kyoto encyclopedia of genes and genomes.** *Nucleic Acids Res* 2000, **28**:27-30.

doi: 10.1186/1471-2164-11-246

**Cite this article as:** Xu *et al.*, Discovery and comparative profiling of microRNAs in a sweet orange red-flesh mutant and its wild type *BMC Genomics* 2010, **11**:246

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

