**Insilico profiling of microRNAs in Korean ginseng (Panax ginseng Meyer)**

Ramya Mathiyalagan, Sathiymoorthy Subramaniam, Sathishkumar Natarajan, Yeon Ju Kim, Myung Suk Sun, Se Young Kim, Yu-Jin Kim and Deok Chun Yang*

Korean Ginseng Center and Ginseng Resource Bank, Kyung Hee University, Yongin 449-701, Korea

Edited by Man-Hee Rhee, Kyungpook National University, Korea

MicroRNAs (miRNAs) are a class of recently discovered non-coding small RNA molecules, on average approximately 21 nucleotides in length, which underlie numerous important biological roles in gene regulation in various organisms. The miRNA database (release 18) has 18,226 miRNAs, which have been deposited from different species. Although miRNAs have been identified and validated in many plant species, no studies have been reported on discovering miRNAs in *Panax ginseng* Meyer, which is a traditionally known medicinal plant in oriental medicine, also known as Korean ginseng. It has triterpene ginseng saponins called ginsenosides, which are responsible for its various pharmacological activities. Predicting conserved miRNAs by homology-based analysis with available expressed sequence tag (EST) sequences can be powerful, if the species lacks whole genome sequence information. In this study by using the EST based computational approach, 69 conserved miRNAs belonging to 44 miRNA families were identified in Korean ginseng. The digital gene expression patterns of predicted conserved miRNAs were analyzed by deep sequencing using small RNA sequences of flower buds, leaves, and lateral roots. We have found that many of the identified miRNAs showed tissue specific expressions. Using the insilico method, 346 potential targets were identified for the predicted 69 conserved miRNAs by searching the ginseng EST database, and the predicted targets were mainly involved in secondary metabolic processes, responses to biotic and abiotic stress, and transcription regulator activities, as well as a variety of other metabolic processes.

**Keywords:** *Panax ginseng*, MicroRNA, Expressed sequence tag, Deep sequencing

**INTRODUCTION**

MicroRNAs (miRNAs) are a class of small, non-protein-coding RNAs with lengths of approximately 21 nucleotides (nt) that act as post-transcriptional regulators in eukaryotes [1]. Like other genes, mature miRNAs also have their own miRNA genes, which are transcribed from their own miRNA genes. In plants, miRNA genes are initially transcribed into primary miRNAs (pri-miRNAs) by pol II [2]. Pri-miRNAs are processed into miRNA precursors (pre-miRNAs) by DICER-LIKE1, which are able to fold into a perfect or near-perfect secondary hairpin structure, and processed into a miRNA duplex (miRNA:miRNA*). It further leads to the release of mature miRNA by the unwinding of the duplexes [1]. Mature miRNAs are assembled into the RNA-induced silencing complex (RISC) to direct the RISC to their complementary target sites in the messenger RNA (mRNA). The activity of miRNA on a target mRNA is dependent on the degree of base pairing, and in the case of perfect or near-perfect base pairing, it leads to target mRNA degradation in plants [3]. Therefore, the perfect
or near-perfect base matching of miRNA to the targets makes the computational prediction of miRNAs easier in plants compared to animals, and they have been successfully applied in many plants [4-6]. miRNA genes are an important class of fine-tuning regulators, playing an important role in a wide range of developmental, biological, and metabolic processes in plants, including metabolism, stress response, vegetative phase change, organogenesis, and signal transduction [7,8].

To date, different approaches have been employed to identify miRNAs in various species, including: 1) direct cloning after isolation of small RNAs with a computational strategy 2) expressed sequence tags (ESTs) analysis, and 3) high throughput sequencing of small RNA [9,10]. Among these three methods, we employed EST analysis and high throughput sequencing of small RNAs to discover Panax ginseng Meyer miRNAs. Comparisons of miRNA of different plant species show that miRNAs have been highly conserved throughout evolution. Its conserved nature helps to identify the miRNA from different plant species by comparative EST based homolog searches, which have been successfully applied in many species, including potato [11], citrus [12], switch grass [13], lettuce [14], and tobacco [15], and it is applicable for those species in which whole genome sequence information is not available [16]. Even though the miRNAs are conserved, some of the miRNAs often express at low levels, or are expressed only in specific tissue or under specific conditions. A new generation of sequencing technologies like high-throughput pyrosequencing technology allows for the identification of lowly expressed or tissue specific expressed miRNA, which was reported in several species such as grapevine [9], tomato [17], and grapevine flower and berry [18].

Based on the annotation criteria, to date 18,226 miRNAs have been deposited in the miRNA registry database (miRBase; release 18.0, http://microrna.sanger.ac.uk) from various species. Although miRNAs have been identified and validated in many plant species, they are largely unknown in P. ginseng (Korean ginseng), which is a traditionally known medicinal plant in oriental medicine where the roots of the plant are mainly used for medicinal purposes. The genus Panax is derived from panacea, which means a cure-all and longevity. It is a slow growing perennial herb of the Araliaceae family, and because of its mysterious power in oriental medicine, people have been using ginseng roots and its extracts to increase physical strength and vigor, and revitalize the body and mind [19]. Ginseng has been used in Korea, as well as other countries such as China and Japan. It contains triterpenoid ginseng saponins called ginsenosides, which are responsible for its various pharmacological activities, including immune system modulation, anti-stress activities, anti-hyperglycemic activities, anti-inflammatory, anti-oxidant, and anti-cancer effects. It also has polysaccharides, flavonoids, peptides, polyacetylic alcohols, and fatty acids [20,21]. In recent years, the increasing evidence of miRNA identification and characterization in other important food crops such as rice, maize, arabidopsis, potato, tomato, citrus, grape fruit, and medicinal tuber crops [22], and also the prediction of terpenoid pathway genes targeting miRNAs in various plant species [11,23,24], as well as evidence of root development related miRNA [25], all induces insight into the analysis of miRNA in P. ginseng. Here, we first report the profiling of miRNA and their targets in P. ginseng (Korean ginseng).

**MATERIALS AND METHODS**

**Plant materials and small RNA sequencing**

The flower buds, leaves, and roots were collected from 6-year field grown P. ginseng plants in south Korea. Immediately after collection, the samples were stored in liquid nitrogen for further analysis. A small RNA library of three samples was constructed using a TruSeq small RNA sample preparation kit, the concentration of RNA was analyzed using Agilent’s Genome Analyzer IIX (GAIIX). The sequence reads were initially trimmed by removing the adapter sequences and low quality sequences with a phred score below 20. Finally, the small RNA sequence was taken in FASTQ format for further bioinformatics analysis.

**Transcriptome sequences and microRNA registry database sequences**

All known miRNAs of mature plants from different plant species were used as reference miRNA for predicting the conserved miRNA in P. ginseng. Known plant miRNAs (reference miRNAs) from the miRBase database (release 17) [26] were derived from different plant species, including Arabidopsis thaliana, Oryza sativa, Glycine max, Brassica napus, Medicago truncatula, Sorghum bicolor, Zea mays, and Saccharum officinarum, as well as all of the other plant species. The complete transcriptome sequences for P. ginseng were collected from the ginseng EST database (http://www.bioherbs.khu.ac.kr/ggrb) to predict the miRNA for P. ginseng [27].
Expressed sequence tag based conserved microRNA prediction

The small RNA raw sequence reads from Illumina’s GAIIx were converted from the FASTQ to the FASTA format, and then the redundancy sequences were removed using the FASTX tool kit (http://hannonlab.cshl.edu/fastx_toolkit), and the remaining unique sequences were selected for further analysis. Sequences 18 to 27 nt bases long were used to do BLASTN against the Rfam database to remove other small RNAs such as transfer RNA (tRNA), ribosomal RNA (rRNA), small interfering RNA (siRNA), small nuclear RNA (snRNA), and small nucleolar RNAs (snoRNA). The remaining sequences were used for the EST based miRNA prediction with the stranded protocol using the mirCheck tool, as well as a custom-made perl script [28]. Firstly, the sequences were matched using the PatScan algorithm with the miRBase database (release 17) to predict the conserved miRNAs in ginseng. Sequences with ≤3 mismatches were considered to be conserved miRNAs in ginseng. Those conserved miRNAs were taken for further analysis, as described in the workflow (Fig. 1).

Target prediction and functional analysis

Predicted miRNA sequences were subjected to target prediction using the web based server psRNATarget (http://plantgrn.noble.org/psRNATarget). This tool has an option to predict user submitted RNAs vs. user submitted transcripts, and we used that option to predict all of the targets. All of the unique transcripts were taken for the functional annotation using the blast2go functional annotation tool. Transcripts were prepared through the de novo assembly and blasted against the non-redundant database, and then subjected to gene ontology analysis.

RESULTS AND DISCUSSION

Generally, miRNAs can be predicted by analysis of
EST and sequencing of small RNAs. Here, we used both of EST analysis and high through put sequencing methodology for the identification of conserved miRNA in *P. ginseng* (Fig. 1).

**Computational identification of conserved microRNA by expressed sequence tag analysis**

The identification of conserved miRNAs by EST analysis is greatly facilitated by the conserved properties of miRNA families among various plant species [29]. To identify the complete set of conserved miRNAs by computational predictions, the availability of the complete genome sequences is a pre-requisite. If complete genomic sequences are lacking, fragmented data like EST and high-throughput genomic sequences have been used [31]. Using the homology based strategy, lots of conserved miRNAs have been identified in various plant species, including potato (*Solanum tuberosum*) [11], switch grass (*Panicum virgatum*) [13], lettuce (*Lactuca sativa*) [14], and rapeseed (*B. napus*) [32]. We employed the computational based approaches to predict miRNAs in *P. ginseng* with available ginseng EST sources from our lab.

miRNA sequences of various plant species were predicted and deposited in the miRBase database [33]. We used miRBase mature miRNA sequences as a reference sequence to predict miRNAs in ginseng using similarity searches. Mature plant miRNAs from different plant species were downloaded from miRBase, redundant sequences were then removed, and non-redundant unique sequences were blasted with *P. ginseng* EST sequences. Using homology searches, 69 miRNAs belonging to 44 conserved miRNA families were identified after repeated and protein coding sequences were removed (Table 1). The majority of predicted miRNAs included miR414, miR1132, miR1439, miR319, miR482, miR847, miR854, miR1436, and miR2628. The mature miRNA sequences were grouped into same member families based on mature miRNA sequence similarity searches using miRBase. In our predictions, the miR414 family was predicted to have the largest abundance of miRNA members (7 members) (Fig. 2), which was also reported in rice (*O. sativa*) [34], Stevia (*Stevia rebaudiana*) [35], and opium poppy (*Papaver somniferum*) [36], while the highest abundance of the same family was reported in switch grass (11 members) [13].

The second largest representative miRNA family was miR1439, where 6 members were identified in our predictions. Only 3 miRNA members of the miR1439 family were predicted in potato [11], whereas 6 members were predicted in *P. ginseng*. Previously, miR1439 was listed as a new rice miRNA [37] and salt induced miRNA in rice [34], and later it was identified in tobacco [15] and potato. Therefore, the prediction of some plant miRNAs in certain plant species may be responsible for special functions, and be conserved in particular species. Another family, miR319, was predicted with 5 members, while miR482, miR847, and miR854 contained 3 members in each family. Additionally, 2 members were contained in each miR1436, miR2628, and miR396 families. The rest of the families were represented with only one member. miR319, reported for various plant development functions like the regulation of leaf senescence, leaf morphogenesis, and leaf complexity [38], and stress regulation of miR319 was reported in sugarcane [39].

Two miRNA847s were reported in *A. thaliana* [40] and *A. lyrata* [41]. Interestingly, in our study, 3 members of miRNA847 were predicted in *P. ginseng*. Another miRNA, miR1436, was identified in this study, which was reported in barley [42], switch grass [13], and rice [34], while 7 members were identified for the same family in potato [11].

We further analyzed the characteristics of conserved miRNAs to distinguish from other small RNAs (Table 1). The length of mature miRNAs varies from 17 to 24 nt, where the majority of miRNAs are confined to 21 nt, followed by 20 and 19 nt (Fig. 3A). The typical lengths of plant mature miRNA sequences are 21 nt, which are in the highest abundance in ginseng miRNAs, similar to other plant species [13,32]. It was reported that the length of pre-miRNAs in plants ranges from 60 to >400 nt [43,44]. The length of precursor miRNAs in *P. ginseng* varies significantly from 55 to 366 nt; however, the majority of pre-miRNAs are 60 to 139 nt in length (Fig. 3B), which is similar to reports of other plant species [11,13]. Having lower MFE is important for the sequences to form stable secondary loop structures for high thermodynamic stability [30]. In this study, the MFE value of identified *P. ginseng* miRNAs ranged from -20.53 to -99.8 kcal/mol, with an average of -35.78 kcal/mol. This MFE value of pre-miRNAs in the present study is consistent with previous reports [32]. MFEI was a valuable criteria used to distinguish potential miRNAs from other types of RNAs. If the MFEI value of the pre-miRNA was higher than 0.85, that sequence was considered to be a potential miRNA [44]. The average MFEI of the predicted *P. ginseng* miRNAs was 0.851 (Table 1).

**Sequence analysis of small RNAs from deep sequencing**

We used the high throughput Illumina sequencing
### Table 1. Identified homology based conserved miRNA in *Panax ginseng*

| miRNA family | Sequence | Length of mature miRNA | Reference | Precursor EST | Length of precursor (3'5') | Location (kcal/mol) | GC% | MFE (kcal/mol) | MFEI |
|--------------|----------|------------------------|-----------|---------------|----------------------------|---------------------|-----|---------------|-----|
| miR156a      | TTTACGGAGATTGGAGAGGAC | 21 | bmu | contig45577 | 60 | 3' | 46.67 | 30 | 0.64 |
| miR159a      | AUAGCAGUAGAGGAGGCUCCU | 21 | esa | contig47841 | 94 | 3' | 44.68 | 31.91 | 0.71 |
| miR164       | UGGAGAAUCAAGGCCCCUGAG | 21 | osa | contig50068 | 248 | 3' | 41.53 | 33.06 | 0.8 |
| miR169h      | GACAUGAGAGAGGAGCAGG | 21 | mtr | contig14564 | 133 | 5' | 29.32 | 20.53 | 0.7 |
| miR172f      | GUAAUCAUGAGACUAGCGCU | 20 | sbi | contig47560 | 287 | 3' | 42.51 | 33.59 | 0.79 |
| miR319b      | UUGGAAGUGGAGAAGCUCCA | 20 | mtr | contig41016 | 72 | 5' | 36.11 | 34.03 | 0.94 |
| miR319e      | UUGGAAGUGGAGAAGCUCCA | 21 | vvi | contig41016 | 72 | 5' | 36.11 | 34.03 | 0.94 |
| miR319f      | UAGCAGUAGAGGCAGCUCCU | 20 | ptc | contig47841 | 94 | 3' | 44.68 | 31.91 | 0.71 |
| miR319g      | UAGGACUGAGGCGAGCCUCU | 20 | ptc | contig54185 | 55 | 3' | 54.55 | 49.45 | 0.91 |
| miR319h      | UAGGACUGAGGCGAGCCUCU | 21 | vvi | contig54185 | 55 | 3' | 54.55 | 49.45 | 0.91 |
| miR396b      | UUCACACAUACUACUACACU | 20 | vvi | contig23595 | 366 | 5' | 33.61 | 25.05 | 0.75 |
| miR396c      | UCCCCGCUUUCUUCUCCCU | 21 | vvi | contig58969 | 263 | 5' | 55.51 | 39.35 | 0.71 |
| miR397       | UCAUUGACAAUUGUUGCUG | 21 | zmu | contig39102 | 88 | 5' | 39.77 | 31.82 | 0.8 |
| miR408a      | AUGCACUCCGUCCUCCCUCGGC | 21 | ath | contig30298 | 102 | 3' | 51.96 | 45.5 | 0.85 |
| miR414d      | UCAUCAUCAUACUACUACU | 21 | ath | contig45947 | 60 | 3' | 45.45 | 42.83 | 0.95 |
| miR414e      | UCAUCAUCAUACUACUACU | 21 | ath | contig45947 | 60 | 3' | 45.45 | 42.83 | 0.95 |
| miR414f      | UCAUCAUCAUACUACUACU | 21 | osa | contig46717 | 328 | 5' | 47.87 | 32.99 | 0.69 |
| miR414h      | UCAUCAUCAUACUACUACU | 21 | osa | contig27881 | 229 | 5' | 41.92 | 32.97 | 0.79 |
| miR414i      | UCAUGGGCUACUACUUCUGA | 21 | ath | contig54204 | 85 | 3' | 44.71 | 35.65 | 0.8 |
| miR414j      | UCAUGGUAACUACUACUACU | 21 | ath | contig27681 | 241 | 5' | 41.49 | 34.44 | 0.83 |
| miR414k      | UCUUGCUACUACUACUACU | 21 | osa | contig55217 | 118 | 5' | 47.46 | 38.39 | 0.81 |
| miR417       | GACAAAAAGAAGAUUGUACA | 21 | ath | contig47470 | 60 | 5' | 28.33 | 37 | 1.31 |
| miR419e      | UUAUGUGAAGAGGAGAAGUGA | 21 | ath | contig58582 | 192 | 3' | 25.52 | 25.78 | 1.01 |
| miR446       | CAUCAUUGAUAGAUCUAGGC | 24 | osa | contig48939 | 106 | 5' | 36.79 | 31.13 | 0.85 |
| miR482       | CUUCUCUCAUCUCCCAUAC | 22 | vvi | contig29669 | 101 | 3' | 46.53 | 41.7 | 0.88 |
| miR482a      | CUCAUCCUCCUCUCCAUCC | 17 | ptc | contig29669 | 101 | 3' | 46.53 | 41.7 | 0.88 |
| miR482c      | CCUUCCUCAUCUCCCAUAC | 20 | ptc | contig29669 | 97 | 3' | 48.45 | 40.41 | 0.83 |
| miR530a      | GCCAUUGCCACUGAUCUUCU | 20 | ptc | contig48663 | 353 | 5' | 42.78 | 30.31 | 0.71 |
| miR783       | AAGAGCUUUCUUCUUCUUCU | 21 | ath | contig18217 | 346 | 5' | 45.38 | 32.95 | 0.73 |
| miR815       | GAGGGAAAGAGGGUAGUGG | 21 | osa | contig59498 | 187 | 3' | 52.94 | 37.33 | 0.71 |
| miR816a      | UUGAUCACUCUACUACUGC | 20 | osa | contig48862 | 90 | 5' | 36.67 | 25.46 | 0.7 |
| miR834b      | UUGUAGUAGUGCGGGUGCCAA | 21 | ath | contig32379 | 67 | 3' | 52.24 | 37.61 | 0.72 |
| miR846       | UUGAUUUUGUGGGGUGAUAU | 21 | ath | contig33367 | 129 | 3' | 34.11 | 27.05 | 0.79 |
| miR847a      | UCAACUCUCUCUCUUCUUGA | 21 | ath | contig50301 | 186 | 5' | 43.01 | 30 | 0.7 |
| miR847b      | UCAACUCUCUCUCUUCUUCUG | 21 | ath | contig46552 | 202 | 5' | 42.57 | 29.06 | 0.68 |
| miR847c      | UCCUUUCUCUCUUCUUCUUAA | 21 | ath | contig00227 | 166 | 3' | 39.76 | 29.4 | 0.74 |
| miR854b      | GAGGAGGAGGAGGAGGAGGAGG | 21 | ath | contig45937 | 119 | 5' | 30.25 | 24.12 | 0.8 |
| miR854d      | GAUGAGAGGAGGAGGAGGAGGAU | 21 | ath | contig33518 | 280 | 5' | 40.71 | 29.05 | 0.71 |
| miR854e      | GAAGAGGAGGAUGAAGAGGAGG | 21 | ath | contig26917 | 169 | 3' | 48.52 | 32.78 | 0.68 |
| miR854f      | GAAGAGGAGGAUGAAGAGGAGG | 21 | ath | contig34488 | 261 | 5' | 42.15 | 29.04 | 0.69 |
| miR1132b     | UAUAUGUGGAGGAGGUGAG | 19 | tae | contig63186 | 273 | 3' | 30.4 | 28.54 | 0.94 |
| miR1134      | CAAACAGAAGAAGAAGAAGAAGA | 24 | tae | contig12044 | 144 | 5' | 25.69 | 31.66 | 0.85 |
Table 1. (Continued)

| miRNA family | Sequence | Length of mature miRNA | Reference | Precursor EST | Length of precursor | Location (3'5') | GC% | MFE (kcal/mol) | MFEI |
|--------------|----------|------------------------|-----------|--------------|-------------------|----------------|-----|----------------|------|
| miR1436c     | UUAUCCUGGGACCGGAGGAGGUU | 21        | osa       | contig61393  | 264               | 3'             | 34.09 | 34.33          | 1.01 |
| miR1436d     | UUAUUAUGGGACCGGAGGAGGUU | 21        | osa       | contig63186  | 275               | 3'             | 30.9  | 80.51          | 0.94 |
| miR1439a     | UUAUGGAUUGGACCGGAGGAGUUAUU | 21       | osa       | contig186765 | 277               | 3'             | 29.96 | 25.13          | 0.84 |
| miR1439b     | UUAUGGAACCGGAGGAGGAGUACU | 21       | osa       | contig56537  | 267               | 3'             | 32.21 | 28.13          | 0.87 |
| miR1439c     | UUAUGGAAUGGACCGGAGGAGUAAU | 21      | osa       | contig56367  | 281               | 3'             | 28.83 | 27.94          | 0.97 |
| miR1439d     | UUAUGGAAUGGACCGGAGGAGUAAU | 21     | osa       | contig10661  | 236               | 3'             | 25    | 22.2           | 0.89 |
| miR1439e     | UUAUGGAAUGGACCGGAGGAGUAAU | 21   | osa       | contig27854  | 283               | 3'             | 29.68 | 23.92          | 0.81 |
| miR1439f     | UUAUGGAAUGGACCGGAGGAGUAAU | 21  | osa       | contig56537  | 283               | 3'             | 29.68 | 23.92          | 0.81 |
| miR1448      | UUAUUCUUAUCUCUCACUCACAC  | 20       | ptc       | contig29669  | 99                | 3'             | 47.47 | 40.91          | 0.87 |
| miR1534a     | UUAUUCUUAUCUCUCACUCACAC  | 20       | gma       | contig49029  | 70                | 3'             | 31.43 | 26.86          | 0.85 |
| miR1886c     | UUAUUCUUAUCUCUCACUCACAC  | 21       | ath       | contig11222  | 98                | 3'             | 42.86 | 38.47          | 0.9  |
| miR20975p    | AGGGAGGGAGGGAGGAAGGAGGAGG | 22     | osa       | contig15753  | 69                | 5'             | 42.03 | 43.62          | 1.04 |
| miR21015p    | AGGGAGGGAGGGAGGAAGGAGGAGG | 22    | osa       | contig17137  | 123               | 5'             | 48.21 | 48.62          | 1.27 |
| miR2108b     | AGGGAGGGAGGGAGGAAGGAGGAGG | 21    | gma       | contig50952  | 65                | 3'             | 32.31 | 46.15          | 1.43 |
| miR2109c     | AGGGAGGGAGGGAGGAAGGAGGAGG | 20    | gma       | contig56006  | 346               | 5'             | 51.45 | 40.64          | 0.79 |
| miR2112b     | AGGGAGGGAGGGAGGAAGGAGGAGG | 21    | ath       | contig55266  | 270               | 3'             | 32.59 | 23.8           | 0.73 |
| miR2606a     | UUAUUCUUAUCUCUCACUCACAC  | 21     | mtr       | contig46524  | 141               | 5'             | 37.59 | 26.88          | 0.72 |
| miR2607      | UUAUUCUUAUCUCUCACUCACAC  | 21     | mtr       | contig38869  | 116               | 5'             | 25.86 | 26.81          | 1.04 |
| miR2626      | UUAUUCUUAUCUCUCACUCACAC  | 21     | mtr       | contig62419  | 56                | 5'             | 39.29 | 50.89          | 1.3  |
| miR2628a     | UUAUUCUUAUCUCUCACUCACAC  | 20     | mtr       | contig23672  | 71                | 5'             | 28.17 | 27.75          | 0.98 |
| miR2628b     | UUAUUCUUAUCUCUCACUCACAC  | 20     | mtr       | contig11869  | 189               | 5'             | 42.33 | 31.64          | 0.75 |
| miR2642      | UUAUUCUUAUCUCUCACUCACAC  | 21     | mtr       | contig70593  | 77                | 5'             | 40.26 | 30.26          | 0.75 |
| miR2643b     | UUAUUCUUAUCUCUCACUCACAC  | 21     | mtr       | contig22496  | 363               | 5'             | 36.08 | 99.8           | 0.76 |
| miR2658      | UUAUUCUUAUCUCUCACUCACAC  | 20     | mtr       | contig28456  | 74                | 3'             | 32.43 | 31.89          | 0.98 |
| miR2665      | UUAUUCUUAUCUCUCACUCACAC  | 21     | mtr       | contig49532  | 60                | 5'             | 33.33 | 27             | 0.81 |
| miR2673b     | UUAUUCUUAUCUCUCACUCACAC  | 22     | mtr       | contig52872  | 189               | 5'             | 55.03 | 41.33          | 0.75 |
| miR2937      | UUAUUCUUAUCUCUCACUCACAC  | 21     | ath       | contig45835  | 79                | 3'             | 43.04 | 41.9           | 0.97 |

miRNA, microRNA; EST, expressed sequence tag; GC, guanine-cytosine content; MFE, minimal free-folding energy; MFEI, minimal free-folding energy index.

Fig. 2. Abundance or frequency of microRNA (miRNA) families in *Panax ginseng*. miR414, miR1439 and miR319 families has highest abundance of miRNAs.

http://dx.doi.org/10.5142/jgr.2013.37.227
technology to sequence small RNAs in *P. ginseng* in order to validate the expression patterns of the EST based predicted conserved *P. ginseng* miRNAs. In high throughput sequencing technology, total of 56,430,729 raw sequences were obtained from the 6-year-old flower buds, leaves, and lateral roots of *P. ginseng*. After removing the low quality sequences, the remaining sequences with length ranging from 17 to 27 nt were obtained. The sequences were further processed to remove other RNAs and redundant sequences. Finally, a total of 5,353,559 non-redundant sequence reads were used for miRNA analysis (Table 2).

**Digital gene expressions of conserved microRNAs in *Panax ginseng* by deep sequencing**

Non-redundant small RNA sequences were used to analyze the digital gene expression pattern of already predicted conserved miRNAs in *P. ginseng*. The small RNA sequences with 100% miRNA sequence similarity with homology based predicted miRNA sequences were used for digital gene expression studies in three tissues. Among the predicted miRNA families by small RNA analysis, miR414 and miR1439 contained the largest number of miRNA with four members, followed by the miR854 family with 3 members. Other families such as miR1436 and miR482 were represented with 2 members in each family. The remaining families had only one member of miRNA.

The expression level of each of the miRNA families also varied. The miRNA family miR482a showed a very high level of expression (number of reads) with the largest number of reads in each organ, such as 740 reads in the flower buds, 13,510 reads in the leaves, and 178 reads in the lateral roots. Followed by, miRNAs such as miR1132h, miR816a, and miR1436d showing the second largest abundant expression of miRNA reads in all three libraries. The miRNAs miR2626, miR1132f, miR1436d and c, miR1439, miR854c and d, and miR414d, e, and f were predicted with >100 miRNA reads in total for all 3 libraries, whereas miRNAs such as miR1534a, miR2658, miR482c, miR414h, and miR156b were predicted with lower expressions.

Tissue specific expression patterns were also observed, as miR1534a was expressed in lateral roots, but it was not expressed in flower buds and leaves, whereas miR414h and miR2097 were detected in flower buds and leaves tissues, but not in lateral roots. The miRNA families miR1448, miR156b, and miR2673b showed expressions in leaves and lateral roots, but not in the flower buds. miRNAs such as miR2658 and miR482c have shown expressions only in leaves, and not in the other tissues (Table 3). Tissue specific expressions of miRNAs were reported.

**Table 2. Distribution of small RNA reads in sequenced *Panax ginseng* tissues**

| Description                              | Leaves | Flower buds | Lateral roots | Total     |
|------------------------------------------|--------|-------------|---------------|-----------|
| Raw sequences                            | 24258021 | 10211169    | 21961539      | 56430729  |
| Adaptor/quality/length (17-27 nt) trimmed | 22803528 | 7021215     | 16777031      | 46601774  |
| Matching trRNA                            | 919530  | 400059      | 470384        | 1789973   |
| Redundant sequence                       | 13625423 | 2789895     | 4714124       | 21129442  |
| Non-redundant sequence                   | 3849681 | 693590      | 810288        | 5353559   |
| Total non-redundant sequence             |         |             |               | 5353559   |

nt. nucleotides; trRNA, transfer RNA/ribosomal RNA.
in various plant species [9, 24]. Even though the root was considered to be the main functional part in *P. ginseng*, leaves and flower buds were also reported for various ginsenosides. This kind of tissue specific expressions of miRNAs represents an interesting topic for further in-depth analysis.

The size distribution patterns of the identified small RNAs in *P. ginseng* were observed such that the majority of the small RNAs were 21 nt in size, followed by 20 nt, 22 nt, and 19 nt, as in the reports of other plant species, such as grapevine [9] and tomato [45].

### Target prediction

Predicting potential targets of miRNA based on a computational approach were aided by the perfect and near perfect complementary characteristics of miRNA with their target mRNA [46]. In order to understand the putative functions of predicted miRNAs, 346 potential targets were identified for the predicted 69 conserved miRNAs by searching the ginseng EST database. Most of the miRNA targets were predicted (Appendix 1), whereas for some miRNAs such as miR482a, miR816, and miR1132, targets were unable to be predicted, which may be due to the limited number of EST sequences available in the databases. Most of the miRNAs were identified with more than one target, especially the miR414 families identified with 68 targets, the miR854 families with 44 targets, and the miR1439 families with 29 targets, which is consistent with the notion that one miRNA may have many targets [47]. Gene Ontology based functional classification of targets was analyzed for understanding the miRNA-gene regulatory network based on biological process and molecular function. In this study, predicted target functions were classified into biological process, molecular function, and cellular component. The main biological process of miRNA targets which involved in transport, protein modification process, regulation of transcription, response to various biotic and abiotic stimulus, secondary metabolic process, and regulation of gene expression which has important role in ginseng (Fig. 4A).

### Table 3. Digital gene expressions of conserved microRNAs (miRNAs) in *Panax ginseng* by deep sequencing

| miRNA family | Mature miRNA sequence | Sequence length | Flower buds | Leaves | Lateral roots |
|--------------|------------------------|-----------------|-------------|--------|--------------|
| miR1132h     | UAUAUGGGACGGAGGUAG     | 19              | 251         | 1235   | 90           |
| miR1436c     | UUAUCCUGGAGGAGGGAGGAU | 21              | 14          | 97     | 3            |
| miR1436d     | UUAUAUGGGACGGAGGUAG    | 21              | 184         | 933    | 61           |
| miR1439a     | UAUAGGAUAGGGAGGAU     | 21              | 2           | 12     | 1            |
| miR1439b     | UUAAGGAACGGAGGAUCU    | 21              | 23          | 150    | 12           |
| miR1439h     | UUAAGGAACGGAGGAUCU    | 21              | 23          | 150    | 12           |
| miR1439c     | UUAAGGAUAGGGAGGAUAU   | 21              | 3           | 28     | 6            |
| miR1448      | CULUCCUAAUCCUCCCAUAC  | 20              | 0           | 15     | 1            |
| miR1534a     | UAULUUGGAGUAUAUAU     | 20              | 0           | 0      | 2            |
| miR169h      | GAACUGAAGAUCUGAAGG    | 21              | 4           | 16     | 1            |
| miR2097-5p   | AGGGAAGGGAGGAAGGAGAGA | 22              | 1           | 16     | 0            |
| miR2626      | AACGUGGGUUAAGGGUGUC   | 21              | 49          | 554    | 7            |
| miR2658      | AUGUGACUULUUUAUGUGC   | 20              | 0           | 4      | 0            |
| miR2673b     | CCCUCUCUCCUCCUCUCUCGC | 22              | 0           | 27     | 7            |
| miR414d      | UCAUCAUCAUCAUCAUCAU   | 21              | 5           | 130    | 2            |
| miR414e      | UCAUCAUCAUCAUCAUCAU   | 21              | 5           | 130    | 2            |
| miR414f      | UCAUCAUCAUCAUCAUCAU   | 21              | 5           | 130    | 2            |
| miR414h      | UCAUCAUCAUCAUCAUCAUAU | 21              | 1           | 6      | 0            |
| miR482a      | CCAUUCCCCUCCUCUACCU   | 17              | 740         | 13510  | 178          |
| miR482c      | CCAUCCCUCCUCCCUACUA   | 20              | 0           | 6      | 0            |
| miR816a      | GUGACAUACUCUUCAUAC    | 20              | 44          | 1244   | 7            |
| miR854b      | GAGGAGGGAGGAGGAGGAGG  | 21              | 13          | 63     | 17           |
| miR854c      | GAUGAGGGAGGAGGAGGAGG  | 21              | 16          | 123    | 20           |
| miR854d      | GAUGAGGGAGGAGGAGGAGG  | 21              | 11          | 86     | 14           |
The predicted putative target genes not only involved in the transcription factors, but also various physiological processes targeting miRNAs were predicted (Appendix 1). Transcription factors were targeted by the miR1439e, miR2109c, miR414h, miR414i, miR5309, miR847a, and miR854 families. In our study, the miR414 family was identified with the largest number of targets, and this miR414 family was reported to be involved in lateral root development in potato [11]. In addition, miR397 and miR1533 were shown to be involved in lateral root development in potato, which the miR397 family was also predicted in *P. ginseng*, whereas miR1533 was initially identified and later removed from *P. ginseng* miRNAs due to the lower MFEI value.

In the present study, miR156a was predicted which was reported for leaf development, vegetative phase change, flowering, and fruit development by targeting the squamosa promoter binding (SPB) protein like family of transcription factors in other plant species [48]. It was also reported that higher levels of expression of miR156/157 could prolong root growth and development in the tuberous medicinal plant [22], but SPB targeting miR156 was unable to be predicted because of the limitless *P. ginseng* EST sequences. miR319, reportedly playing an important function in leaf morphogenesis [49], was identified in *P. ginseng*.
Ginsenosides, very important triterpenoid secondary metabolites in the medicinal plant *P. ginseng*, were reported for their various pharmacological properties. Genes such as 3-hydroxy-3-methylglutaryl-coenzyme A reductase (HMGR), farnesyl-diphosphate synthase (FPS), geranyl-diphosphate synthase, squalene synthase, and squalene epoxidase (SE) were reported as putative ginsenoside pathway genes [27], and hydroxylation by cytochrome P450 and glycosylation mediated by UDP-glycosyltransferases lead to synthesis of various ginsenosides. Overexpression of *P. ginseng* squalene synthase was shown to increase the ginsenoside production [50]. These putative ginsenoside pathway genes were predicted as the miRNA targets, especially SE targeting miR854b and miR854c. To support this, previous reports have shown that SE was the target of miRNAs, especially miR1533 [11]. Previous reports on the target identification showed that HMGR and FPS were targeted by different miRNAs [11,23]. Accordingly, our results also showed that miR854e was identified to target FPS, while miRNA targeting HMGR was also predicted, but due to the lower MEFI value, it was removed in our analysis. Various cytochrome and glucosyltransferase targeting miRNAs were predicted in this study, as in the reports of other plant species [11,22]. Ginsenoside Ro is the only oleanane-type pentacyclic triterpene, which is a minor component in *P. ginseng*, and has different pharmacological effects. Beta-amyrib synthase converts 2,3-oxidosqualene to beta-aminin, which leads to the production of oleanane type ginsenosides (Ro). miRNAs such as miR1439b and miR1439h were predicted to target beta amyrin synthase in *P. ginseng*, which was also reported in potato [11]. Various reports have shown a high similarity between predicted miRNA and their targets to previously reported miRNA and their targets. Alternatively, our miRNAs and target predictions showed less similarity with previously reported known miRNAs and their targets. The lower availability of *P. ginseng* transcriptomes in GenBank, and the lower number of phylogenetic relations, or the lower similarity with other known crops, could be one of the possible reasons for less conservation in nature of *P. ginseng* miRNAs compared to other known miRNAs.

Some of the conserved miRNAs are expressed lower or below detection level in the case of the number of reads of small RNA sequences analyzed in flower buds, leaves, and lateral root tissues, and it may be present in other tissues that have not yet been analyzed. Most of the miRNA predictions in other plant species mainly used the young stage in their samples, whereas in contrast, we used fully matured tissues to sequence small RNA. These may be possible causes for the less expressed miRNAs in *P. ginseng* analyzed tissues. Numerous ginseng specific novel miRNAs may show a high level of expression in other tissues or organs, or different developmental stages are yet to be investigated and further experiments would provide more species specific miRNAs.

To sum up, we discovered 69 miRNAs in Korean ginseng, and tissue specific expression patterns of the identified miRNAs were analyzed using digital gene expressions of deep sequenced small RNAs of the flower buds, leaves, and lateral roots. Therefore, these results provide a basis for the regulatory roles of miRNA in ginseng. To get better insight into the miRNAs in ginseng, further studies on small RNA sequencing from specific tissues will be carried out.

**ACKNOWLEDGEMENTS**

The research funding was supported by the Korea Institute of Planning & Evaluation for Technology in Food, Agriculture, Forestry & Fisheries (KIPET no. 309019-3) and by a grant from the Next-Generation BioGreen 21 Program (SSAC, grant#: PJ00952903), Rural Development Administration, Republic of Korea. The ginseng sample used in this study was provided by Kyung Hee University, South Korea.

**REFERENCES**

1. Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 2004;116:281-297.
2. Zhou X, Ruan J, Wang G, Zhang W. Characterization and identification of microRNA core promoters in four model species. PLoS Comput Biol 2007;3:e37.
3. Jones-Rhoades MW, Bartel DP, Bartel B. MicroRNAs and their regulatory roles in plants. Annu Rev Plant Biol 2006;57:19-53.
4. Lu Y, Yang X. Computational identification of novel microRNAs and their targets in *Vigna unguiculata*. Comp Funct Genomics 2010;pii: 128297.
5. Sunkar R, Zhu JK. Novel and stress-regulated microRNAs and other small RNAs from *Arabidopsis*. Plant Cell 2004;16:2001-2019.
6. Zhang B, Pan X, Anderson TA. Identification of 188 conserved maize microRNAs and their targets. FEBS Lett 2006;580:3753-3762.
7. Ambros V. The functions of animal microRNAs. Nature 2004;431:350-355.
10. Song C, Wang C, Zhang C, Korir NK, Yu H, Ma Z, Fang J. Deep sequencing discovery of novel and conserved microRNAs in trifoliate orange (Citrus trifoliata). BMC Genomics 2010;11:431.

11. Xie F, Frazier TP, Zhang B. Identification, characterization and expression analysis of microRNAs and their targets in the potato (Solanum tuberosum). Gene 2011;473:8-22.

12. 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.

13. Xie F, Frazier TP, Zhang B. Identification and characterization of microRNAs and their targets in the bioenergy plant switchgrass (Panicum virgatum). Planta 2010;232:417-434.

14. Han Y, Zhu B, Luan F, Zhu H, Shao Y, Chen A, Lu C, Luo Y. Conserved miRNAs and their targets identified in lettuce (Lactuca) by EST analysis. Gene 2010;463:1-7.

15. Frazier TP, Xie F, Freistaedter A, Burklew CE, Zhang B. Identification and characterization of microRNAs and their target genes in tobacco (Nicotiana tabacum). Planta 2010;232:1289-1308.

16. Zhao CZ, Xia H, Frazier TP, Yao YY, Bi YP, Li AQ, Li MJ, Li CS, Zhang BH, Wang XJ. Deep sequencing identifies novel and conserved microRNAs in peanuts (Arachis hypogaea L.). BMC Plant Biol 2010;10:3.

17. Mohorianu I, Schwach F, Jing R, Lopez-Gomollon S, Moxon S, Szitnya G, Sorefan K, Moulton V, Dalmay T. Profiling of short RNAs during fleshy fruit development reveals stage-specific sRNAome expression patterns. Plant J 2011;67:232-246.

18. Wang C, Wang X, Kibet NK, Song C, Zhang C, Li X, Han J, Fang J. Deep sequencing of grapevine flower and berry short RNA library for discovery of novel microRNAs and validation of precise sequences of grapevine microRNAs deposited in miRBase. Physiol Plant 2011;143:64-81.

19. Vogler BK, Pittler MH, Ernst E. The efficacy of ginseng: A systematic review of randomized clinical trials. Eur J Clin Pharmacol 1999;55:567-575.

20. Choi KT. Botanical characteristics, pharmacological effects and medicinal components of Korean Panax ginseng C A Meyer. Acta Pharmacol Sin 2008;29:1109-1118.

21. Kim SK, Park JH. Trends in ginseng research in 2010. J Ginseng Res 2011;35:389-398.

22. Yang Y, Chen X, Chen J, Xu H, Li J, Zhang Z. Differential miRNA expression in Rehmannia glutinosa plants subjected to continuous cropping. BMC Plant Biol 2011;11:53.

23. Pani A, Mahapatra RK, Behera N, Naik PK. Computational identification of sweet wormwood (Artemisia annua) microRNA and their mRNA targets. Genomics Proteomics Bioinformatics 2011;9:200-210.

24. Xu Q, Liu Y, Zhu A, Wu X, Ye J, Yu K, Guo W, Deng X. Discovery and comparative profiling of microRNAs in a sweet orange red-flesh mutant and its wild type. BMC Genomics 2010;11:246.

25. Meng Y, Ma X, Chen D, Wu P, Chen M. MicroRNA-mediated signaling involved in plant root development. Biochem Biophys Res Commun 2010;393:345-349.

26. Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ. miRBase: tools for microRNA genomics. Nucleic Acids Res 2008;36(Database issue):D154-D158.

27. Sathiyaamoorthy S, In JG, Lee BS, Kwon WS, Yang DU, Kim JH, Yang DC. Insilico analysis for expressed sequence tags from embryogenic callus and flower buds of Panax ginseng C. A. Meyer. J Ginseng Res 2011;35:21-30.

28. Jones-Rhoades MW. Prediction of plant miRNA genes. Methods Mol Biol 2010;592:19-30.

29. Zhang B, Pan X, Cannon CH, Cobb GP, Anderson TA. Conservation and divergence of plant microRNA genes. Plant J 2006;46:243-259.

30. Zhang B, Pan X, Stellwag EJ. Identification of soybean microRNAs and their targets. Planta 2008;229:161-182.

31. Sunkar R, Jagadeeswaran G. Insilico identification of conserved microRNAs in large number of diverse plant species. BMC Plant Biol 2008;8:37.

32. Dhandapani V, Ramchiary N, Paul P, Kim J, Choi SH, Lee J, Hur Y, Lim YP. Identification of potential microRNAs and their targets in Brassica rapa L. Mol Cells 2011;32:21-37.

33. Kozomara A, Griffiths-Jones S. miRBase: integrating microRNA annotation and deep-sequencing data. Nucleic Acids Res 2011;39(Database issue):D152-D157.

34. 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.

35. Guleria P, Yadav SK. Identification of miR414 and expression analysis of conserved miRNAs from Stevia rebaudiana. Genomics Proteomics Bioinformatics 2011;9:211-217.
opium poppy (*Papaver somniferum* L.). Plant Cell Rep 2010;29:757-769.

37. Luo YC, Zhou H, Li Y, Chen JY, Yang JH, Chen YQ, Qu LH. Rice embryogenic calli express a unique set of microRNAs, suggesting regulatory roles of microRNAs in plant post-embryogenic development. FEBS Lett 2006;580:5111-5116.

38. Schommer C, Bresso EG, Spinelli S, Palatnik J. Role of microRNA miR319 in plant development. In: Sunkar R, ed. MicroRNAs in plant development and stress responses. Berlin: Springer, 2012. p.29-47.

39. Thiebaut F, Rojas CA, Almeida KL, Grativol C, Domiciano GC, Lamb CR, Engler Jde A, Hemerly AS, Ferreira PC. Regulation of miR319 during cold stress in sugarcane. Plant Cell Environ 2012;35:502-512.

40. 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.

41. Fahlgren N, Jogdeo S, Kasschau KD, Sullivan CM, Chapman EJ, Laubinger S, Smith LM, Dasenko M, Given SA, Weigel D et al. MicroRNA gene evolution in *Arabidopsis lyrata* and *Arabidopsis thaliana*. Plant Cell 2010;22:1074-1089.

42. Kantar M, Unver T, Budak H. Regulation of barley miRNAs upon dehydration stress correlated with target gene expression. Funct Integr Genomics 2010;10:493-507.

43. Smalheiser NR, Torvik VI. Mammalian microRNAs derived from genomic repeats. Trends Genet 2005;21:322-326.

44. Zhang BH, Pan XP, Cox SB, Cobb GP, Anderson TA. Evidence that miRNAs are different from other RNAs. Cell Mol Life Sci 2006;63:246-254.

45. Moxon S, Jing R, Szittya G, Schwach F, Rusholme Pilcher RL, Moulton V, Dalmary T. Deep sequencing of tomato short RNAs identifies microRNAs targeting genes involved in fruit ripening. Genome Res 2008;18:1602-1609.

46. Jones-Rhoades MW, Bartel DP. Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. Mol Cell 2004;14:787-799.

47. Lewis BP, Burge CB, Bartel DP. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell 2005;120:15-20.

48. Wu G, Poethig RS. Temporal regulation of shoot development in *Arabidopsis thaliana* by miR156 and its target SPL3. Development 2006;133:3539-3547.

49. Palatnik JF, Allen E, Wu X, Schommer C, Schwab R, Carrington JC, Weigel D. Control of leaf morphogenesis by microRNAs. Nature 2003;425:257-263.

50. Shim JS, Lee OR, Kim YJ, Lee JH, Kim JH, Jung DY, In JG, Lee BS, Yang DC. Overexpression of *PgSQS1* increases ginsenoside production and negatively affects ginseng growth rate in *Panax ginseng*. J Ginseng Res 2010;34:98-103.
### Appendix 1. Target prediction of identified microRNAs (miRNAs) in *Panax ginseng*

| miRNA family | Target protein                                                                 | Target ID     |
|--------------|--------------------------------------------------------------------------------|---------------|
| miR1134      | 1-O-acylglucose:anthocyanin-O-acyltransferase-like protein                      | Contig60089   |
| miR1134      | NAC domain-containing                                                           | Contig56197   |
| miR1134      | Oligopeptide transporter OPT family                                            | Contig23794   |
| miR1134      | PDI-like protein                                                               | Contig45218   |
| miR1134      | RESA-like protein with                                                           | Contig45096   |
| miR1134      | Ribosome biogenesis protein                                                     | Contig45306   |
| miR1134      | RNA recognition motif-containing protein                                         | Contig61121   |
| miR1436c     | Cytochrome C oxidase polypeptide                                                | Contig48926   |
| miR1436d     | Protein kinase                                                                  | Contig90914   |
| miR1439a     | Chromatin remodeling complex subunit                                            | Contig45689   |
| miR1439a     | Enzyme of the cupin superfamily                                                 | Contig42038   |
| miR1439a     | Proton-dependent oligopeptide transport family protein                          | Contig30763   |
| miR1439a     | Synaptic glycoprotein SC2                                                        | Contig35006   |
| miR1439b     | Beta-amyrin synthase                                                            | Contig54769   |
| miR1439b     | Chromatin remodeling complex subunit                                            | Contig45689   |
| miR1439b     | Disease resistance protein                                                       | Contig11515   |
| miR1439b     | Enzyme of the cupin superfamily                                                 | Contig42038   |
| miR1439b     | Nitroreductase family protein                                                    | Contig53658   |
| miR1439c     | Chromatin remodeling complex subunit                                            | Contig45689   |
| miR1439c     | Disease resistance protein                                                       | Contig11515   |
| miR1439c     | Enzyme of the cupin superfamily                                                 | Contig42038   |
| miR1439c     | Transcription factor jumonji domain-containing protein                          | Contig45164   |
| miR1439d     | Armadillo beta-catenin repeat family protein                                     | Contig45166   |
| miR1439d     | Disease resistance protein                                                       | Contig11515   |
| miR1439d     | Enzyme of the cupin superfamily                                                 | Contig42038   |
| miR1439e     | Enzyme of the cupin superfamily                                                 | Contig42038   |
| miR1439e     | Nucleic acid binding                                                            | Contig46706   |
| miR1439e     | Protein phosphatase                                                             | Contig38335   |
| miR1439e     | Serine endopeptidase                                                            | Contig33916   |
| miR1439e     | Squalene monooxygenase                                                           | Contig47397   |
| miR1439e     | TATA-associated factor II 58                                                    | Contig51124   |
| miR1439e     | WRKY transcription                                                              | Contig59342   |
| miR1439h     | Amino acid                                                                      | Contig46145   |
| miR1439h     | Beta-amyrin synthase                                                            | Contig54769   |
| miR1439h     | Chromatin remodeling complex subunit                                            | Contig45689   |
| miR1439h     | Disease resistance protein                                                       | Contig11515   |
| miR1439h     | Enzyme of the cupin superfamily                                                 | Contig42038   |
| miR1439h     | Nitroreductase family protein                                                    | Contig53658   |
### Appendix 1. (Continued)

| miR1448 | CC-NBS-LRR resistance protein | Contig55112 |
|---------|-------------------------------|-------------|
| miR1448 | Cinnamyl alcohol dehydrogenase-like protein | Contig48716 |
| miR1448 | Fsh11 (protease 11) ATP-dependent peptidase ATPase metallopeptidase | Contig23505 |
| miR1448 | mRNA binding protein precursor | Contig49813 |
| miR1448 | Pentatricopeptide repeat-containing | Contig49516 |
| miR1534a | Acyl- oxidase | Contig09994 |
| miR1534a | Calcium-dependent protein | Contig33156 |
| miR1534a | Chromosome region maintenance protein 1 | Contig48609 |
| miR1534a | Cytochrome c oxidase polypeptide vc | Contig48926 |
| miR1534a | Cytochrome p450 monooxygenase CYP72A59 | Contig11848 |
| miR1534a | Elongation factor 1-alpha | Contig15951 |
| miR1534a | Fat domain-containing protein | Contig49029 |
| miR1534a | Glucan synthase component | Contig49469 |
| miR1534a | Glutathione peroxidase | Contig51675 |
| miR1534a | Lipase class 3 family protein | Contig42396 |
| miR1534a | Lon protease | Contig13984 |
| miR1534a | Multidrug resistance protein | Contig52708 |
| miR1534a | Phosphomethyl pyrimidine kinase thiamin-phosphate pyrophosphorylase | Contig52615 |
| miR1534a | Pre-mRNA splicing factor rna | Contig10498 |
| miR1534a | Pyrophosphate-energized vacuolar membrane proton | Contig46517 |
| miR1534a | R2R3-myb transcription factor myb11 | Contig23436 |
| miR156a | Dead deah box helicase family protein | Contig45577 |
| miR156a | Hypothetical Protein | Contig40276 |
| miR156a | Soluble starch synthase iv-2 | Contig47036 |
| miR156a | Vitamin-b12 independent methionine 5-methyltetrahydropteroyltriglutamate-homocysteine | Contig51422 |
| miR159a | Delta-1-pyrroline-5-carboxylate dehydrogenase | Contig26802 |
| miR1886c | Cation chloride cotransporter | Contig52276 |
| miR1886c | F-box family protein | Contig35247 |
| miR1886c | 26s protease regulatory subunit | Contig24305 |
| miR1886c | Heterogeneous nuclear ribonucleoprotein A2 | Contig33032 |
| miR1886c | Ketose-bisphosphate aldolase class-ii family protein | Contig15415 |
| miR172f | Calcium-binding allergen OLE | Contig33032 |
| miR172f | Lipase class 3 family protein | Contig52276 |
| miR172f | Type ii peroxiredoxin | Contig57487 |
| miR1886c | Cation chloride cotransporter | Contig50840 |
| miR1886c | CCAAT-binding transcription factor family protein | Contig48908 |
| miR1886c | Ketose-bisphosphate aldolase class-ii family protein | Contig11222 |
| miR1886c | Phospholipase D | Contig53739 |
| miR1886c | Receptor protein kinase clavata1 | Contig30888 |
| miRNA       | Target Name                                                                 | Contig     |
|------------|-----------------------------------------------------------------------------|------------|
| miR2097-5p | 2OG-FE oxygenase family protein                                             | Contig20782|
| miR2097-5p | ABA response element binding factor                                          | Contig36126|
| miR2097-5p | Cellulose synthase                                                          | Contig11555|
| miR2097-5p | DNA binding                                                                 | Contig52665|
| miR2097-5p | Gamma-adaptin 1                                                             | Contig58073|
| miR2097-5p | Heat shock                                                                  | Contig15069|
| miR2097-5p | MCA1 (mid1-complementing activity 1)                                         | Contig45919|
| miR2097-5p | Nucleolar protein                                                           | Contig15060|
| miR2097-5p | Trehalose-6-phosphate synthase                                              | Contig27435|
| miR2101-5p | ELP1 (edm2-like protein1)                                                   | Contig32343|
| miR2101-5p | Inositol-tetrakisphosphate 1                                                | Contig46141|
| miR2101-5p | Meprin and traf homology domain-containing protein math domain-containing protein | Contig45263|
| miR2101-5p | Protein binding                                                             | Contig48832|
| miR2101-5p | S-adenosylmethionine-dependent methyltransferase                            | Contig52110|
| miR2101-5p | SPL1-related2 protein                                                       | Contig17762|
| miR2108b   | Serine threonine protein kinase                                              | Contig48746|
| miR2108b   | With no lysine kinase                                                       | Contig45290|
| miR2109c   | Transcription factor, putative                                               | Contig43484|
| miR2112b   | C2 domain-containing protein                                                | Contig44978|
| miR2112b   | Cytochrome                                                                  | Contig46603|
| miR2112b   | Transcriptional repressor                                                  | Contig23975|
| miR2606a   | ARF1-binding protein                                                        | Contig31431|
| miR2606a   | ATP binding                                                                 | Contig51306|
| miR2606a   | Heat shock protein 70 -interacting                                          | Contig35223|
| miR2607    | Cytosolic phosphoglucomutase                                                | Contig21451|
| miR2607    | Potassium transporter                                                       | Contig10118|
| miR2626    | Obtusifoliol 14-alpha demethylase                                           | Contig46095|
| miR2626    | Zinc finger (C3HC4-type ring finger) family protein                          | Contig50055|
| miR2626a   | Bromodomain protein                                                         | Contig30826|
| miR2628b   | mRNA splicing                                                               | Contig45688|
| miR2642    | Cinnamoyl- reductase                                                        | Contig49468|
| miR2642    | Cytochrome c6                                                               | Contig62852|
| miR2642    | Exocyst complex subunit SEC15-like family protein                           | Contig51982|
| miR2642    | Pectinacetyltransferase family protein                                       | Contig16753|
| miR2642    | Photosystem i PSAH protein                                                  | Contig57701|
| miR2642    | Plasma membrane h+ -ATPase                                                  | Contig19168|
| miR2642    | Serine-threonine protein plant                                              | Contig51145|
| miR2643b   | TPR repeat-containing protein                                               | Contig48906|
| miR2658    | Homeodomain leucine zipper protein                                          | Contig28456|
| miR2658    | Metalloendopeptidase                                                        | Contig51850|
| miR2658    | Phosphoinositide binding                                                    | Contig62161|
| miR2665    | 3-phosphoserine phosphatase                                                 | Contig49532|
### Appendix 1. (Continued)

| miR   | Description                                                                 | Contig |
|-------|------------------------------------------------------------------------------|--------|
| miR2665 | Ap2 ERF domain-containing transcription factor                                | 56891  |
| miR2665 | Diacylglycerol acyltransferase                                                 | 48735  |
| miR2665 | Multidrug resistance protein ABC transporter family                           | 17769  |
| miR2673b | 2-cys peroxiredoxin                                                          | 49523  |
| miR2673b | 6b-interacting protein 1                                                      | 37386  |
| miR2673b | CBS domain-containing protein                                                  | 13027  |
| miR2673b | Della protein                                                                 | 45937  |
| miR2673b | E3 ubiquitin ligase                                                           | 47870  |
| miR2673b | Glycine-rich protein 2b                                                       | 38426  |
| miR2673b | Glycine-rich RNA-binding protein                                               | 60922  |
| miR2673b | H Aca ribonucleoprotein complex subunit 1-like protein 1                       | 54326  |
| miR2673b | Inositol phosphate kinase                                                      | 51804  |
| miR2673b | Kinesin light                                                                | 55627  |
| miR2673b | Phospholipid cytidylytransferase                                              | 32791  |
| miR2673b | Protein kinase                                                                | 46112  |
| miR2673b | Ribosomal protein 117-like protein                                             | 53111  |
| miR2673b | Hypothetical protein                                                          | 49580  |
| miR2673b | Tata-binding protein-associated factor 2n-like                                 | 53856  |
| miR2673b | WRKY transcription                                                            | 62618  |
| miR2673b | Zinc finger                                                                  | 37605  |
| miR2937 | Dme DNA n-glycosylase DNA-(apurinic or apyrimidinic site) lyase                | 45773  |
| miR2937 | Dynamin-related protein expressed                                             | 47241  |
| miR2937 | Heat shock protein binding protein                                             | 23728  |
| miR2937 | Phospho ribosylformylglycinamidine synthase                                   | 56549  |
| miR2937 | Serine-threonine protein plant                                                | 48306  |
| miR319b | Transcription factor WRKY4                                                    | 48636  |
| miR319b | L1 specific homeobox gene atml1 ovule-specific homeobox protein a20           | 50022  |
| miR319b | Receptor protein kinase clavata1                                              | 30156  |
| miR319e | Transcription factor WRKY4                                                    | 48636  |
| miR319e | L1 specific homeobox gene atml1 ovule-specific homeobox protein a20           | 50022  |
| miR319e | Receptor protein kinase clavata1                                              | 30156  |
| miR319f | Delta-1-pyrroline-5-carboxylate dehydrogenase                                | 26802  |
| miR319f | Proteasome subunit alpha type 3                                              | 25577  |
| miR319g | Five finger-containing phosphoinositide                                        | 1379   |
| miR319g | Phototropic-responsive NPH3 family protein                                    | 36246  |
| miR319h | Ubiquitin-protein PUB49                                                       | 8309   |
| miR319h | Phototropic-responsive NPH3 family protein                                    | 36246  |
| miR319h | Stromal membrane-associated                                                   | 19616  |
| miR396b | Acyl- oxidase                                                                | 99994  |
| miR396b | Beta-glucosidase-like protein                                                 | 97198  |
| miR396b | Heat shock protein                                                            | 53352  |
| miR396c | ABC transporter family protein                                                | 48065  |
| miRNA   | Gene Name                      | Contig       |
|---------|-------------------------------|--------------|
| miR396c | AP2 ERF domain-containing transcription factor | Contig24365 |
| miR396c | ELF3 homologue                | Contig10723  |
| miR396c | Mitochondrial substrate carrier | Contig46804  |
| miR396c | Splicing factor                | Contig58523  |
| miR396c | Type-B response regulator      | Contig33927  |
| miR397  | Actin                         | Contig20734  |
| miR397  | Cell division protein          | Contig47714  |
| miR397  | Cytosolic malate dehydrogenase | Contig39102  |
| miR397  | Multidrug resistance-associated protein | Contig23354 |
| miR397  | Synaptic glycoprotein SC2      | Contig35004  |
| miR408a | Chemocyanin precursor          | Contig57069  |
| miR414d | 60S ribosomal protein l6       | Contig51433  |
| miR414d | ADP-glucose pyrophosphorylase family protein | Contig52871 |
| miR414d | Ascorbate peroxidase           | Contig36806  |
| miR414d | CBL-interacting serine threonine-protein | Contig46717 |
| miR414d | Conserved hypothetical protein | Contig46471  |
| miR414d | Cytochrome p450                | Contig61265  |
| miR414d | Late embryogenesis abundant protein LEA14 | Contig49375 |
| miR414d | NLI interacting factor family protein | Contig12161 |
| miR414d | Pre-mRNA-splicing factor CWC-22 | Contig30561 |
| miR414d | Protein phosphatase            | Contig39881  |
| miR414d | Ring finger containing         | Contig48197  |
| miR414d | RNA helicase                   | Contig12914  |
| miR414e | 60S ribosomal protein l6       | Contig51433  |
| miR414e | ADP-glucose pyrophosphorylase family protein | Contig52871 |
| miR414e | Ascorbate peroxidase           | Contig36806  |
| miR414e | CBL-interacting serine threonine-protein | Contig46717 |
| miR414e | Conserved hypothetical protein | Contig46471  |
| miR414e | Cytochrome p450                | Contig61265  |
| miR414e | Late embryogenesis abundant protein LEA14 | Contig49375 |
| miR414e | NLI interacting factor family protein | Contig12161 |
| miR414e | Pre-mRNA-splicing factor CWC-22 | Contig30561 |
| miR414e | Protein phosphatase            | Contig39881  |
| miR414e | Ring finger containing         | Contig48197  |
| miR414e | RNA helicase                   | Contig12914  |
| miR414f | 60S ribosomal protein l6       | Contig51433  |
| miR414f | ADP-glucose pyrophosphorylase family protein | Contig52871 |
| miR414f | Ascorbate peroxidase           | Contig36806  |
| miR414f | CBL-interacting serine threonine-protein | Contig46717 |
| miR414f | Conserved hypothetical protein | Contig46471  |
| miR414f | Cytochrome p450                | Contig61265  |
| miR414f | Late embryogenesis abundant protein LEA14 | Contig49375 |
| miR414f | NLI interacting factor family protein | Contig12161 |
### Appendix 1. (Continued)

| miRNA | Description | Contig |
|-------|-------------|--------|
| miR414f | Pre-mRNA-splicing factor CWC-22 | Contig30561 |
| miR414f | Protein phosphatase | Contig39881 |
| miR414f | Ring finger containing | Contig48197 |
| miR414f | RNA helicase | Contig12914 |
| miR414h | 60s ribosomal protein l6 | Contig51433 |
| miR414h | Ascorbate peroxidase | Contig36806 |
| miR414h | CBL-interacting serine threonine-protein | Contig46717 |
| miR414h | Conserved hypothetical protein | Contig66471 |
| miR414h | Pre-mRNA-splicing factor Cwc-22 | Contig30561 |
| miR414h | Ring finger containing | Contig48197 |
| miR414i | Luminal binding protein | Contig16316 |
| miR414i | Pentatricopeptide repeat-containing protein | Contig48382 |
| miR414i | Zip transporter | Contig21400 |
| miR414j | Cytochrome P450 reductase | Contig45338 |
| miR414j | Heat shock factor | Contig50711 |
| miR414j | Kinase family protein | Contig24024 |
| miR414j | Lectin protein kinase family protein | Contig21894 |
| miR414j | MYB transcription factor | Contig16861 |
| miR414j | Phosphatidylinositol-4-phosphate 5-kinase family protein | Contig50328 |
| miR414j | Small RAS-like GTP-binding protein | Contig67528 |
| miR414j | Tryptophanyl-tRNA synthetase | Contig34813 |
| miR414j | Vascular morphogenesis protein | Contig68171 |
| miR414l | Copalyl diphosphate synthase | Contig52365 |
| miR414l | DNA binding | Contig36982 |
| miR414l | Heat shock factor protein HSF30 | Contig23567 |
| miR414l | Insulinase containing expressed | Contig30114 |
| miR414l | Leucine-rich repeat-containing | Contig54865 |
| miR414l | Pescadillo-like protein | Contig30996 |
| miR414l | RNA polymerase ii transcription elongation factor SPT5 | Contig31454 |
| miR414l | Ubiquitin-protein ligase 1 | Contig17692 |
| miR417 | Binding protein | Contig37222 |
| miR417 | Nucleic acid binding | Contig33218 |
| miR419e | Auxin response factor 4 | Contig30161 |
| miR419e | Beta-glucosidase 8 | Contig50325 |
| miR419e | Bromodomain protein | Contig45057 |
| miR419e | Bzip transcription factor | Contig52353 |
| miRNA | Description                                                                                     | Contig  |
|-------|------------------------------------------------------------------------------------------------|--------|
| miR419e | Dna binding protein                                   | Contig49741    |
| miR419e | Heavy-metal-associated domain-containing protein                         | Contig61082    |
| miR419e | Nucleosome assembly                                   | Contig33375    |
| miR419e | Polyphenol oxidase                                     | Contig27778    |
| miR419e | SIT4 phosphatase-associated family protein             | Contig27966    |
| miR446 | Beta-galactosidase like protein                        | Contig45051    |
| miR482 | Alpha-glucosidase                                      | Contig47628    |
| miR530a | COP1-interacting protein 7                           | Contig30564    |
| miR530a | ISP4-like protein                                     | Contig45556    |
| miR530a | Zinc finger protein                                   | Contig27700    |
| miR783 | Pentatricopeptide repeat-containing                   | Contig55799    |
| miR783 | Short-chain dehydrogenase reductase family protein     | Contig52837    |
| miR783 | Vacuolar protein sorting-associated                    | Contig50413    |
| miR815 | Adenylate kinase                                      | Contig34836    |
| miR815 | Chromatin remodeling complex subunit                   | Contig43242    |
| miR815 | Cullin-like 1 protein                                 | Contig30323    |
| miR815 | Dead-box protein                                      | Contig31488    |
| miR815 | Fat domain-containing protein                          | Contig29906    |
| miR815 | Glutamyl-tRNA amidotransferase subunit A               | Contig34203    |
| miR815 | Lysosomal alpha-glucosidase                           | Contig49082    |
| miR815 | RPH1 (resistance to phytophthora 1)                    | Contig38523    |
| miR834b | Histone H3                                            | Contig50215    |
| miR834b | Receptor-like serine threonine protein kinase ARK3     | Contig45114    |
| miR834b | Set domain protein                                    | Contig46763    |
| miR846 | Glucan endo beta-glucosidase                          | Contig42642    |
| miR846 | Kip-related cyclin-dependent kinase inhibitor 7        | Contig47511    |
| miR846 | Multidrug resistance protein                          | Contig49696    |
| miR846 | RNA-binding protein CP31                              | Contig49296    |
| miR847a | Bile acid: sodium symporter family protein            | Contig58467    |
| miR847a | Heat shock protein                                    | Contig01966    |
| miR847a | Viral A-type inclusion protein                        | Contig47741    |
| miR847a | Zinc finger                                           | Contig51073    |
| miR847b | Amino acid binding                                    | Contig46296    |
| miR847b | Amino acid permease                                   | Contig18864    |
| miR847b | Binding protein                                       | Contig30978    |
| miR847b | Chlorophyll a, b-binding protein                      | Contig51038    |
| miR847b | Gerany/geranyl pyrophosphate synthase-related protein | Contig49005    |
| miR847b | Inositol -trisphosphate 5 6 kinase                    | Contig46827    |
| miR847b | Serine threonine protein kinase                       | Contig34614    |
| miR847b | TPR domain containing protein                        | Contig45092    |
| miR847b | Vitamin-B12 independent methionine 5-methyltetrahydrofolateglutamate-homocysteine | Contig23325    |
| miR847c | 2-dehydro-3-deoxyphosphoheptonate aldolase 3-deoxy-d-arabino-heptulosonate 7-phosphate synthetase | Contig28077    |
### Appendix 1. (Continued)

| miR847c | Amidohydrolase domain-containing protein | Contig56377 |
| miR847c | Cytosolic phosphoglycerate kinase 1 | Contig13961 |
| miR847c | DNA repair protein RAD4 family | Contig36417 |
| miR847c | Eukaryotic translation initiation factor 4g | Contig17702 |
| miR847c | PSI type ii chlorophyll a b-binding protein | Contig52383 |
| miR847c | Small nuclear ribonucleoprotein E | Contig56309 |
| miR847c | Vacuolar ATP synthase subunit | Contig47586 |
| miR847c | Vq motif-containing protein | Contig55306 |
| miR847c | Zinc finger | Contig48641 |
| miR854b | CRR3 (chlororespiratory reduction 3) | Contig33843 |
| miR854b | Ethylene responsive element binding factor | Contig49258 |
| miR854b | MYB-related transcription factor LBM2-like | Contig59010 |
| miR854b | Squalene epoxidase | Contig52768 |
| miR854b | Starch branching enzyme ii | Contig30265 |
| miR854b | Transcription initiation factor iib | Contig34296 |
| miR854b | UDP-N-acetylglucosamine: dolichol phosphate N-acetylglucosamine-1-p transferase | Contig32764 |
| miR854c | Transcription factor WRKY4 | Contig56799 |
| miR854c | CRR3 (chlororespiratory reduction 3) | Contig53843 |
| miR854c | DNA binding | Contig35357 |
| miR854c | Ethylene responsive element binding factor | Contig49258 |
| miR854c | F-box family protein | Contig47190 |
| miR854c | Gamma response i protein | Contig21322 |
| miR854c | Hypothetical protein | Contig56684 |
| miR854c | Phototrophic-responsive NPH3 family protein | Contig46596 |
| miR854c | Plastid division protein | Contig55829 |
| miR854c | Polynucleotide phosphorylase | Contig32431 |
| miR854c | Protein binding protein | Contig34488 |
| miR854c | RNA helicase | Contig12919 |
| miR854c | SAS10 U3 ribonucleoprotein family protein | Contig27731 |
| miR854c | Squalene epoxidase | Contig52768 |
| miR854c | Starch branching enzyme ii | Contig30265 |
| miR854c | Transcription initiation factor iib | Contig34296 |
| miR854c | Type i phosphodiesterase nucleotide pyrophosphatase family protein | Contig6494 |
| miR854c | U4 u6 small nuclear ribonucleoprotein PRP3 | Contig26841 |
| miR854c | Ubiquitin-conjugating enzyme e2 I | Contig14762 |
| miR854d | Transcription factor WRKY4 | Contig56799 |
| miR854d | Aminocyl-tRNA synthetase family | Contig15841 |
| miR854d | C-4 sterol methyl oxidase | Contig48724 |
| miR854d | Chalcone isomerase | Contig52143 |
| miR854d | Galactosyltransferase family protein | Contig4535 |
| miR854d | Heat shock protein 90 | Contig6778 |
| miR854d | PDV2 (plastid division2) | Contig53900 |
### Appendix 1. (Continued)

| miRNA   | Gene Name                                      | Contig  |
|---------|-----------------------------------------------|---------|
| miR854d | Serine threonine protein kinase                | Contig21634 |
| miR854e | Cell division protein                          | Contig11097 |
| miR854e | Farnesylation diphosphate synthase              | Contig50044 |
| miR854e | Glutathione reductase                          | Contig12638 |
| miR854e | Inositol-tetrakisphosphate 1                   | Contig33622 |
| miR854e | Multicatalytic endopeptidase proteasome beta subunit | Contig48772 |
| miR854e | Pf068 protein                                  | Contig53517  |
| miR854e | Phospholipase D                                | Contig51301  |
| miR854e | Pre-mRNA splicing factor PRP38                 | Contig48032  |
| miR854e | TCP family transcription                       | Contig47975   |
| miR854e | Type I phosphodiesterase nucleotide pyrophosphatase family protein | Contig46494  |