



Comprehensive Annotation and Functional Exploration of MicroRNAs in Lettuce

OPEN ACCESS

Edited by:

Lei Wang,
Institute of Botany, Chinese Academy
of Sciences (CAS), China

Reviewed by:

Jianbo Xie,
Beijing Forestry University, China
Yun Ju Kim,
Institute for Basic Science (IBS),
South Korea

*Correspondence:

Xinyu Guo
guoxy@nercita.org.cn
Lei Li
lei.li@pku.edu.cn
Xiaozeng Yang
yangxiaozeng@baafs.net.cn;
yangxz@smaworld.com

† These authors have contributed
equally to this work and share first
authorship

Specialty section:

This article was submitted to
Plant Physiology,
a section of the journal
Frontiers in Plant Science

Received: 23 September 2021

Accepted: 28 October 2021

Published: 24 December 2021

Citation:

Deng Y, Qin Y, Yang P, Du J,
Kuang Z, Zhao Y, Wang Y, Li D, Wei J,
Guo X, Li L and Yang X (2021)
Comprehensive Annotation
and Functional Exploration
of MicroRNAs in Lettuce.
Front. Plant Sci. 12:781836.
doi: 10.3389/fpls.2021.781836

Yang Deng^{1,2†}, Yajuan Qin^{1,2†}, Pan Yang^{1,3,4†}, Jianjun Du^{1,5†}, Zheng Kuang^{1,4},
Yongxin Zhao^{1,2}, Ying Wang⁴, Dayong Li^{1,6}, Jianhua Wei^{1,2}, Xinyu Guo^{1,5*}, Lei Li^{4*} and
Xiaozeng Yang^{1,2*}

¹ Beijing Academy of Agriculture and Forestry Sciences, Beijing, China, ² Beijing Key Laboratory of Agricultural Genetic Resources and Biotechnology, Beijing Agro-Biotechnology Research Center, Beijing, China, ³ College of Life Sciences, Fujian Agriculture and Forestry University, Fuzhou, China, ⁴ State Key Laboratory of Protein and Plant Gene Research, Peking-Tsinghua Center for Life Sciences, School of Advanced Agricultural Sciences and School of Life Sciences, Peking University, Beijing, China, ⁵ Beijing Key Lab of Digital Plant, Beijing Research Center for Information Technology in Agriculture, Beijing, China, ⁶ Beijing Key Laboratory of Vegetable Germplasm Improvement, Beijing Vegetable Research Center, Beijing, China

MicroRNA (miRNA) is an important endogenous post-transcriptional regulator, while lettuce (*Lactuca sativa*) is a leafy vegetable of global economic significance. However, there are few studies on miRNAs in lettuce, and research on miRNA regulatory network in lettuce is absent. In this study, through deep sequencing of small RNAs in different tissues, together with a reference genome, 157 high-confidence miRNA loci in lettuce were comprehensively identified, and their expression patterns were determined. Using a combination of computational prediction and high-throughput experimental verification, a set of reliable lettuce miRNA targets were obtained. Furthermore, through RNA-Seq, the expression profiles of these targets and a comprehensive view of the negative regulatory relationship between miRNAs and their targets was acquired based on a correlation analysis. To further understand miRNA functions, a miRNA regulatory network was constructed, with miRNAs at the core and combining transcription factors and miRNA target genes. This regulatory network, mainly composed of feed forward loop motifs, greatly increases understanding of the potential functions of miRNAs, and many unknown potential regulatory links were discovered. Finally, considering its specific expression pattern, *Lsa-MIR408* as a hub gene was employed to illustrate the function of the regulatory network, and genetic experiments revealed its ability to increase the fresh weight and achene size of lettuce. In short, this work lays a solid foundation for the study of miRNA functions and regulatory networks in lettuce.

Keywords: lettuce, microRNA (miRNA), sRNA-Seq, expression correlation, regulatory network, miR408

INTRODUCTION

Lettuce (*Lactuca sativa*), of the Asteraceae family, is a leafy vegetable with important economic value. According to the Food and Agriculture Organization¹, the global production of lettuce, together with chicory, exceeded 29 million tons in 2019, ranking third of all leafy vegetables. In the United States alone, lettuce production in 2019 exceeded 3.6 million tons, with a market value of over 1.5 billion USD. The worldwide popularity of lettuce is due to its diverse cultivars, taste, and high nutritional value, being rich in vitamins, folates, and minerals (Tamura et al., 2018). With the availability of a reference genome (Reyes-Chin-Wo et al., 2017) and numerous transcriptomes (Zhang L. et al., 2017), lettuce has become a model for studying Asteraceae and related plants. Lettuce has numerous cultivars, including butterhead, crisphead, looseleaf, romaine, stem, and many mixed types, which are highly varied in morphology, and provide good material for studying the morphological development of plants (Seki et al., 2020; Yu et al., 2020). The same is true for lettuce in terms of leaf colors (Su et al., 2020), metabolites (Zhang et al., 2020), heat responses (Chen et al., 2017, 2018), viral resistance (Inderbitzin et al., 2019), etc. Lettuce is also an important example of crop domestication and for adaptation research (Zhang L. et al., 2017; Wei et al., 2021). Recently, lettuce has been used as a novel platform for biopharmaceutical production due to its numerous advantages, including low processing cost, consistent and scalable production, and the excellent biosafety profile of transgenic plants (Daniell et al., 2019; Zhang et al., 2019).

MicroRNA (miRNA) is an important post-transcriptional regulatory factor, which has received considerable research interest (Fromm et al., 2020). In plants, major mature miRNAs only have 20–24 nucleotides (nts), which can find target mRNAs through nearly perfect sequence complementation, leading to degradation or translation block (Voinnet, 2009). MiRNA genes share many common features with protein-coding genes such as defined promoters and transcription start sites as well as the exon-intron gene structure (Voinnet, 2009). In general, RNA polymerase II (Pol II) is responsible for transcriptional activation of miRNA genes in plants (Lee et al., 2004). The initially transcribed miRNA, called pri-miRNA, has a stem-loop structure (Bartel, 2009). In plants, pri-miRNA is first cut into pre-miRNA by Dicer-like, and then further processed into a duplex composed of mature miRNA and miRNA*. Only the mature miRNA guides the miRNA-induced silencing complex to regulate its mRNA targets, and other parts are degraded (Bartel, 2009; Voinnet, 2009). Studies have found that plant miRNAs have very diverse functions in development, responses to environmental challenges, etc. (Chandra et al., 2017). For example, miR156 can regulate the transition from juvenile to adult developmental stages (Wang et al., 2009; Wu et al., 2009), miR319/159 is related to leaf morphology and reproduction (Yang et al., 2013; Cheng et al., 2020), and miR399 is associated with the phosphate-starvation response (Fujii et al., 2005). It has recently been discovered that miRNAs in plants can also be used as carriers

for information exchange between different species (Shahid et al., 2018; Tsikou et al., 2018).

In the last decade, high-throughput sequencing has become the most powerful method to identify miRNAs. The combination of small RNA sequencing (sRNA-Seq) and subsequent bioinformatic analysis has uncovered a large number of new miRNAs (Guo et al., 2020). The method of miRNA target gene exploration is constantly improving, which further improved the prediction accuracy of miRNA targets as well (Zhao et al., 2021). The high-throughput target gene verification method, parallel analysis of RNA ends sequencing [PARE-Seq (German et al., 2009)] or degradome sequencing enables identification of miRNA target genes on a genomic scale. In addition, combined with RNA-Seq data, the correlation of expression patterns between miRNAs and potential targets provides not only a reference for target gene prediction, but also the potential for the discovery of new regulatory miRNA-target pairs. Motif scanning, together with methods such as Chromatin Immunoprecipitation Sequencing (Johnson et al., 2007) or DNA Affinity Purification Sequencing (Bartlett et al., 2017), has been widely employed as a high-throughput means to detect the regulation between transcriptional factors (TFs) and their targets, which could be used to discover the upstream regulation from TFs to miRNAs. With all these technical advantages, a whole genome scaled picture of miRNA-based regulatory networks in a specific species can be achieved. For example, a multiple layer regulation network was recently constructed in *Arabidopsis* (Gao et al., 2021). However, despite being an important economic crop, case studies of miRNAs in lettuce are scarce (Huo et al., 2016), and the systematic identification and annotation of miRNAs and their regulatory networks are absent.

In this study, by sequencing sRNAs in four different tissues and using a pipeline centered by miRDeep-P2 (Kuang et al., 2019), 157 high-confidence miRNAs were identified in lettuce. The detailed characteristics of these miRNAs were annotated, including sequences, structure, conservation, clusters, selection after duplication, and expression patterns in different tissues. The targets of these miRNAs were predicted by three different methods, and Gene Ontology [GO (Mi et al., 2019)] and Kyoto Encyclopedia of Genes and Genomes [KEGG (Ogata et al., 1999)] enrichment analyses indicated that these target genes are widely involved in signal transduction, protein binding, and various enzyme activities, partially reflecting the diversification of miRNA functions and the role of a specific group of miRNA targets, TFs. Furthermore, a correlation analysis on miRNA and target expression patterns was performed, and the negative regulatory relationship between miRNAs and target genes is highlighted, although this correlation is affected by numerous factors. In addition, a regulatory network involving miRNAs, TFs, and targets, was constructed, and the connection between different feed forward loops (FFLs) helped in defining a number of new regulatory links. Finally, based on its specific expression pattern and potential function, miR408 was used as an example to verify this exploration of miRNAs and miRNA regulatory network. Genetic experiments not only confirmed the solidity of the data, but also found that *Lsa-MIR408* can promote the vegetative growth of lettuce and increase achene size. This

¹<http://www.fao.org/>

work systematically identified and annotated miRNAs in lettuce, and conducted preliminary explorations of the functions of these miRNAs by means of target genes, expression correlation, regulatory networks, etc., and lays a solid foundation for further research on lettuce miRNAs.

MATERIALS AND METHODS

Plant Materials

Lettuce (*Lactuca sativa* L.) achenes were purchased from the vegetable institute of Beijing Academy of Agriculture and Forestry Sciences, Beijing. Achenes were first sown on a wet filter paper in a petri dish and moved directly to the soil in plastic pots after germination. The seedlings were cultured in a greenhouse (16 h/8 h day/night cycle, light intensity $200 \mu\text{mol m}^{-2} \text{s}^{-1}$, relative humidity 30–50%). The root, stem, and leaf were collected from 4-week-old plants. The flowers were collected at full bloomed capitula (19-week-old plants). After collection, the samples of each tissue were quickly frozen in liquid nitrogen and stored at -80°C for the extraction of total RNA, respectively.

Small RNA and mRNA Libraries Construction and Sequencing

The RNA from roots, stems, leaves, and flowers of lettuce was, respectively, isolated using the OminiPlant RNA Kit (Cwbio, China), following the manufacturer's protocol. The integrity and quality of RNAs were validated by an Agilent 2100 Bioanalyzer. Small RNA cDNA libraries were prepared using the Small RNA Sample Prep Kit (Illumina, United States), and mRNA cDNA libraries were prepared using the TruSeq Stranded RNA LT Kit (Illumina, United States), based on the manufacturer's protocols. Briefly, sRNAs were isolated from 20 μg of total RNA by 15% polyacrylamide gel electrophoresis and ligated two adaptors, including a 5'-RNA and a 3'-RNA adaptor. Then, the samples were converted and amplified to cDNA by RT-PCR (mRNA cDNA library preparation is similar). Lastly, the validated cDNA libraries were sequenced by Illumina HiSeq2500. Small RNA and mRNA cDNA libraries included two biological replicates, and 16 libraries were produced.

Identification of Conserved, Asteraceae-Specific, and Lettuce-Specific MicroRNAs

After sRNA library sequencing, clean reads were obtained by filtering low-quality sequences, including junk reads, adaptor sequences, polyA tags and reads <18 bp and >30 bp. The extracted clean reads (19–25 nt in length) were used for miRNA prediction. Reads matching plant non-coding RNAs, including tRNA, rRNA snRNA, and snoRNA sequences in the Rfam database (version 13.0; Kalvari et al., 2018), with no more than one mismatch were further filtered. Next, the remaining sequences were mapped to the lettuce genome (Reyes-Chin-Wo et al., 2017), and candidate miRNAs were detected via miRDeep-P2 software (version 1.1.4; Kuang et al., 2019). The adjacent sequences of mapped sRNAs were extracted as candidate miRNA

precursor sequences (details in the miRDeep-P2 manual). The secondary structures of all candidates were predicted by RNAfold (version 2.1.2; Tav et al., 2016). The newly updated plant miRNA criteria (Axtell and Meyers, 2018) were employed to identify miRNAs.

To annotate the conservation of miRNAs in lettuce, all the predicted mature miRNA sequences with ± 1 nt adjacent nucleotide were aligned with all plant miRNAs in PmiREN1.0 (Guo et al., 2020), with no more than two mismatches, using Bowtie (version 1.2.2; Langmead, 2010) software. Those with no matched miRNAs in PmiREN1.0 were assigned as lettuce-specific miRNAs, while those which only hit miRNAs in other Asteraceae species were considered Asteraceae-specific miRNAs, and others were annotated as conserved.

Synteny Analysis of Identified MicroRNAs

The synteny analysis of miRNAs was carried out using JCVI (MCscanX python version 1.1.18; Wang et al., 2012). Firstly, the transformed Browser Extensible Data (BED) file and protein sequences were set as input files, and the collinearity blocks across the whole lettuce genome were obtained. Subsequently, the genomic locations of all miRNAs were reflected to these collinearity blocks via BEDTools (version 2.26.0; Quinlan and Hall, 2010; subfunction: intersect), and members from the same miRNA family that followed similar gene orders in these collinearity blocks were considered as syntenic miRNAs. The result plot was generated with Circos (version 0.69.9; Krzywinski et al., 2009).

Targets Prediction of Identified MicroRNAs

Two plant miRNA target prediction toolkits, psRNATarget (Dai et al., 2018) and RNAhybrid (Kruger and Rehmsmeier, 2006), were used to predict miRNA targets. The identified lettuce miRNA sequences and mRNA transcript sequences were uploaded to psRNATarget webserver² and the latest default parameters (2017 release) were used. Targets with an E -value ≤ 3.0 were kept as possible miRNA targets. Meanwhile, RNAhybrid (version 2.1.2) was used to predict energetically plausible miRNA-mRNA duplexes with plant-specific constraints where “ $-d$ ” parameter was set as “8,12.” A strict cut-off value for minimum free energy/minimum duplex energy of 0.75 was used (Alves et al., 2009).

Degradome-seq (PARE-seq) data downloaded from GEO datasets (SRP078275) were also used to predict miRNA targets. CleaveLand4.0 (Addo-Quaye et al., 2009) software was used to identify putative miRNA cleavage sites with default settings. The reads were aligned to lettuce transcript sequences to generated density files. The predicted miRNA mature sequences were aligned to transcript sequences to identify potential miRNA target sites. The density distribution of reads and miRNA-mRNA alignment were used together to classify miRNA target candidates. All potential miRNA targets were assigned into one

²<https://www.zhaolab.org/psRNATarget/>

of five categories, and to reduce false positives, only results from categories 0, 1, and 2 were retained.

MicroRNA and mRNA Expression Analysis

For miRNA expression, expression of mature and star miRNAs was normalized by RPM. For each sRNA-seq dataset, reads mapped to pri-miRNAs (no mismatches allowed) and localized in genomic positions of mature miRNAs (no more than 2-nt shift allowed) were considered to correspond to mature miRNAs. The total numbers of these reads were counted to calculate the RPM for the mature miRNAs.

For mRNA expression, all raw mRNA reads were checked by Fastqc (version 0.11.9; Brown et al., 2017) and low-quality reads and adapters were removed by Cutadapt (version 3.2; Kechin et al., 2017). Clean reads were aligned to the lettuce genome using HISTAT2 (version 2.1.0; Kim et al., 2019). StringTie (version 2.1.5; Pertea et al., 2015) was used to assemble the transcript and quantify each mRNA-seq dataset. The expression values of transcripts were normalized by fragments per kilobase per million fragments (FPKM) and the results of StringTie were extracted.

Correlation Analysis Between the Expression of MicroRNAs and Predicted Targets

The Pearson correlation coefficient between the expression matrix of miRNAs and targets was calculated using the R function “cor.test()”. Firstly, according to the biological replicates of four tissues, the raw expression matrix of miRNAs and targets were averaged and transposed by the R function “t()” to obtain a matrix with miRNAs and targets as factors. Then, the “for loop” of R language, was used to calculate the correlation coefficient between each miRNA and target, and a Student’s *t*-test was used to test significance. The results were displayed by the R package (Pheatmap), where blue bars showed positive correlation, red bars showed negative correlation, and a *P*-value of <0.05 is presented as “*” (Figure 3A).

Network Construction

For the miRNA network, TFs of lettuce were annotated by searching all mRNA transcripts of lettuce against TF datasets from PlantTFDB³ (Jin et al., 2017), and the default filtered result was kept. Upstream 2000 nts were extracted as putative promoter sequences for each pre-miRNAs and miRNA target transcripts, and then submitted into PlantRegMap webserver⁴ (Tian et al., 2020) to scan TF binding sites. FFL motifs were selected by an in-house perl script based on these interactions between TFs-miRNAs, miRNAs-targets and TF-miRNA targets. Then, a genome-wide miRNA network was constructed in terms of all FFL motifs and Cytoscape (version 3.7.1; Shannon et al., 2003) was employed to display the network structure (Figure 5B).

³<http://planttfdb.gao-lab.org/prediction.php>

⁴<http://plantregmap.gao-lab.org/>

Construction of Plant Expression Vector and Lettuce Transformation

To construct the miR408 overexpression vector, genomic DNA of lettuce containing pre-miR408 was PCR amplified using the primers MIR408-F/R (Supplementary Table 22). This fragment was then cloned into pBI121 vector after *Xba*I and *Sac*I double digestion, to create the MIR408-overexpressing vector 35S:MIR408. The vector containing the overexpression cassette for MIR408 was then introduced into the *Agrobacterium tumefaciens* strain EHA105 by the freeze-thaw method. Cotyledon explants of lettuce plants were transformed with *Agrobacterium* EHA105 containing the 35S:MIR408 construct, following the reported protocol.

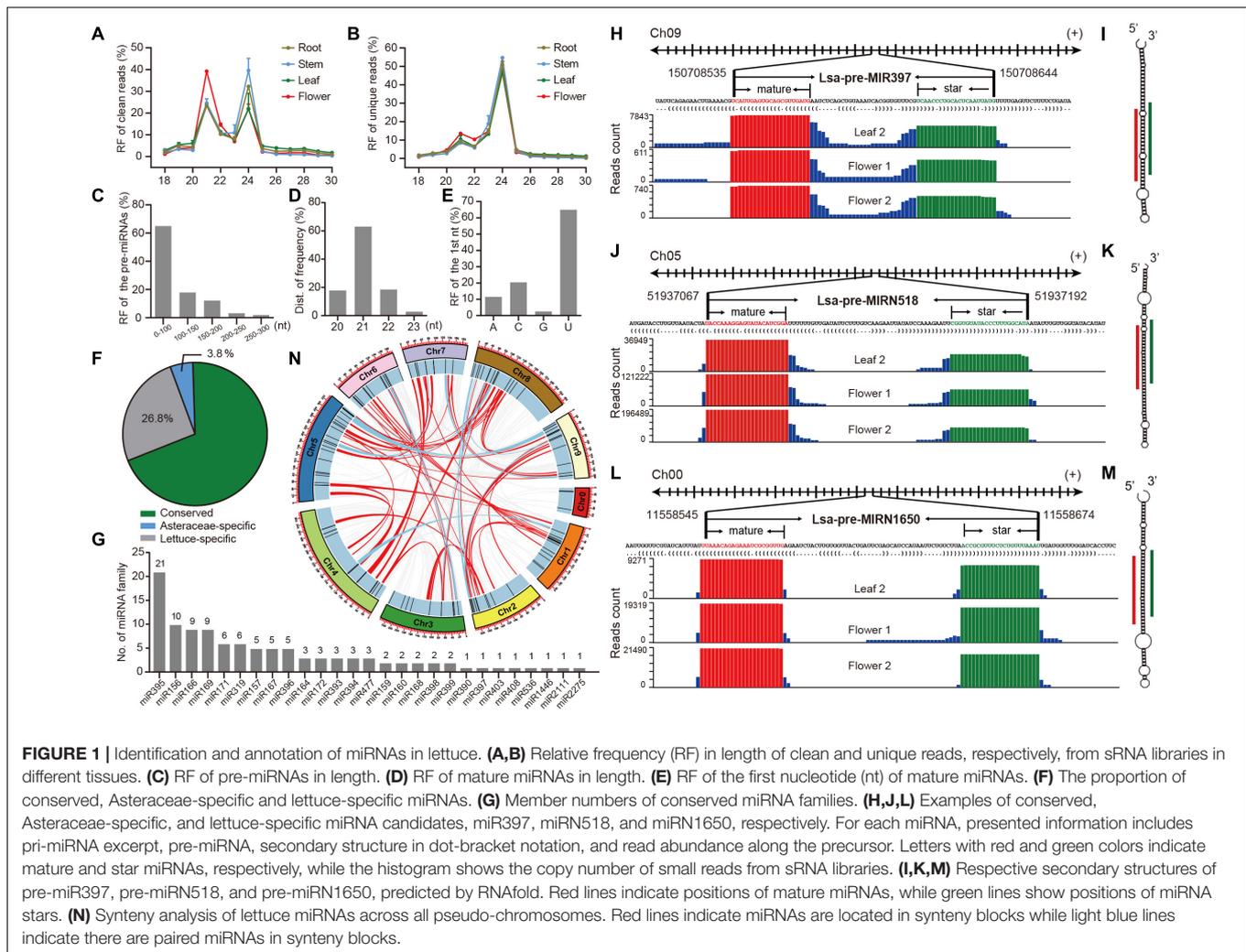
Quantitative RT-PCR

Total RNA from lettuce was reverse transcribed using the Maxima First Strand cDNA Synthesis Kit (ThermoFisher, United States). The resultant cDNA was analyzed by SYBR premix Ex Taq (ABMgood, United States) with the CFX96 Touch System (Bio-Rad, United States). The PCR protocol was 95°C for 30 s, 40 cycles of 95°C for 1 min, and 60°C for 10 s. 18S ribosomal RNA was used as the internal control gene. Total small RNA was extracted by the miRcute miRNA Isolation Kit (TianGen, China), following the manufacturer’s instructions. Reverse miRNA transcription was conducted using the miRcute Plus miRNA First-Strand cDNA SynthesisKit (TianGen, China). The resultant cDNA was analyzed using the miRcute miRNA qPCR Detection Kit, SYBR Green (TianGen, China), with the CFX96 Touch System (Bio-Rad, United States). PCR was performed under the following conditions: 95°C for 15 min, 34 cycles of 94°C for 30 s, and 60°C for 34 s. U6 snRNA was used as the reference gene. PCR reactions were carried out with three biological replicates and each biological replicate was performed with four technical repeats. The results were calculated using the $2^{-\Delta\Delta Ct}$ method. All primers used for qRT-PCR analysis are listed in Supplementary Table 22.

RESULTS

Comprehensive Identification and Annotation of MicroRNAs in Lettuce

To comprehensively identify miRNAs in lettuce, sRNA libraries from different tissues, including leaf, stem, root and flower, were prepared in duplicate (details in section “Materials and Methods”) and observed with well consistency (Supplementary Figure 1A). These eight libraries produced 240 million raw reads and 235 million clean reads (Supplementary Table 1). Length distribution of these sRNA libraries showed two peaks at 21 and 24 nt in both clean and unique sRNAs, potentially reflecting the abundance of miRNAs and small interference RNAs (siRNAs), respectively (Figures 1A,B and Supplementary Tables 2, 3). The large decrease at 21 nt from clean to unique read distributions further supports this speculation (Figures 1A,B) because, in general, miRNAs prefer to accumulate many identical reads, while siRNAs possess much greater



sequence diversity (Carthew and Sontheimer, 2009; Axtell and Meyers, 2018). Compared to other tissues, abundance at the 21-nt peak is higher than that at 24-nt in the flower (Figure 1A), suggesting higher miRNA activities in flower tissue.

MiRDeep-P2 (Kuang et al., 2019) was employed to identify miRNAs when only 19–25 nt reads were selected as inputs. Filtered with strict plant miRNA criteria (Axtell and Meyers, 2018), 157 miRNA loci, belonging to 67 families, were annotated (Supplementary Table 4). Several features of the miRNA loci were scanned, and most of their precursors, accounting for 80%, were less than 150 nts (Figure 1C), and the length of mature miRNAs were predominantly 21 nts (Figure 1D), while the first nucleic acid of mature miRNAs is frequently U (Figure 1E). All of these observations meet plant miRNA characteristics, as indicated by previous research (Axtell and Meyers, 2018), suggesting that the identification and annotation is reliable. A conservation search against all miRNA entries in PmiREN1.0 (Guo et al., 2020), found that 109, including 27 families, are conserved since their counterparts could be found in other non-Asteraceae species, while only six are considered Asteraceae-specific, because counterparts

were only detected in Asteraceae species. Forty-two items were annotated as lettuce specific, considering no conserved ones were detected (Supplementary Table 4 and Figure 1F). Among the conserved families, miR395 has the most abundant members, while the other 18 families have multiple members (Figure 1G). The annotation for each miRNA includes all relevant information, including genomic location, sequences (Supplementary Table 4), secondary structure, and reads signature along precursors. Figures 1H–M, show conserved, Asteraceae-specific and lettuce-specific examples, respectively. Standard stem-loop structure and reads signature, such as the accumulation at mature and star miRNAs, strongly support them being highly confident candidates.

The miRNA clusters were further scanned, and nine clusters were obtained when a cluster region length not exceeding 10 kb was defined (Supplementary Table 5). There were four miR395 clusters, containing 13 miR395 members (61.2% of 21 miR395s), indicating why miR395 is the most abundant miRNA family, and its family expansion may be due to segmental duplication after a tandem duplication, considering all four clusters are located at the same chromosome (Supplementary Table 5 and

Figure 1G). Since there exists a whole genome triplication (WGT) event during the evolution of the lettuce genome (Reyes-Chin-Wo et al., 2017), the WGT impact on miRNA selection was further investigated. Interestingly, even though there are clearly numerous synteny blocks (**Supplementary Tables 6, 7**), only six were kept in two-paired synteny blocks after the WGT event (**Supplementary Table 7**); most duplicated miRNA members were lost during the genome evolution (**Figure 1N**).

Expression Patterns of Lettuce MicroRNAs in Different Tissues

The replicated tissue samples enabled a systematic examination of the expression patterns of lettuce miRNAs. Normalized by reads per million (RPM), the complete expression profile of 157 miRNAs was obtained (**Figure 2A**, **Supplementary Table 8**, and **Supplementary Figure 1B**). In general, conserved miRNAs had higher average expression (**Figure 2B**). MiRNAs varied in their expression patterns, and some were consistently highly expressed in different tissues, while some varied significantly between different tissues (**Figure 2A**). When the low (<10 RPM) and high expression (>500 RPM) was defined, approximately 16 miRNAs were consistently highly expressed in different tissues, such as house-keeping genes, which account for 10% of all miRNAs (**Figures 2C,D**), while 31 miRNAs were highly expressed in a specific tissue (**Figures 2C,E**). Over half of the miRNAs were expressed at an ordinary level, and approximately 14% were had constantly low expression (**Figure 2C**). To further explore the expression pattern of miRNAs in different tissues, a *k*-means cluster analysis was performed. When *k* = 16, the expression patterns of all miRNAs could be separated well (**Figure 2F**, **Supplementary Figure 2**, and **Supplementary Table 8**).

Differently expressed miRNAs were compared among the sampled tissues. In general, 35 to 64 miRNAs were up or down-regulated between a pair of tissues (**Figures 2G,H**, **Supplementary Figure 3A**, and **Supplementary Table 9**), but less than 10 of them were differently expressed in multiple tissues (**Figure 2I**). Analysis of a specific miRNA or miRNA family identified that, as in other species, miR166 is the house-keeping miRNA, highly expressed in all examined tissues (**Figure 2D**). As expected, considering their roles in the transition between juvenile and adult, miR156 was highly expressed in juvenile leaf tissue, while miR172 was highly expressed in flower tissue (Wang et al., 2009; Wu et al., 2009; **Figure 2E**). However, Lsa-miR156i, an miR156 member, was highly expressed in flower tissue (**Figure 2E**). Meanwhile, a number of lettuce-specific miRNAs were consistently highly expressed in all tissues, including Lsa-miRN1709, Lsa-miRN1652a/b, and Lsa-miRN1664 (**Figure 2D**), and some of them, including Lsa-miRN1, Lsa-miRN40, Lsa-miRN69 (**Figure 2E**), were highly expressed in a specific tissue. This was less common, as species-specific miRNAs, in general, had low expression (Yang et al., 2021).

Integrative Target Identification of Lettuce MicroRNAs

To elucidate the function of miRNAs in lettuce, a combination of computational methods and high-throughput experiments

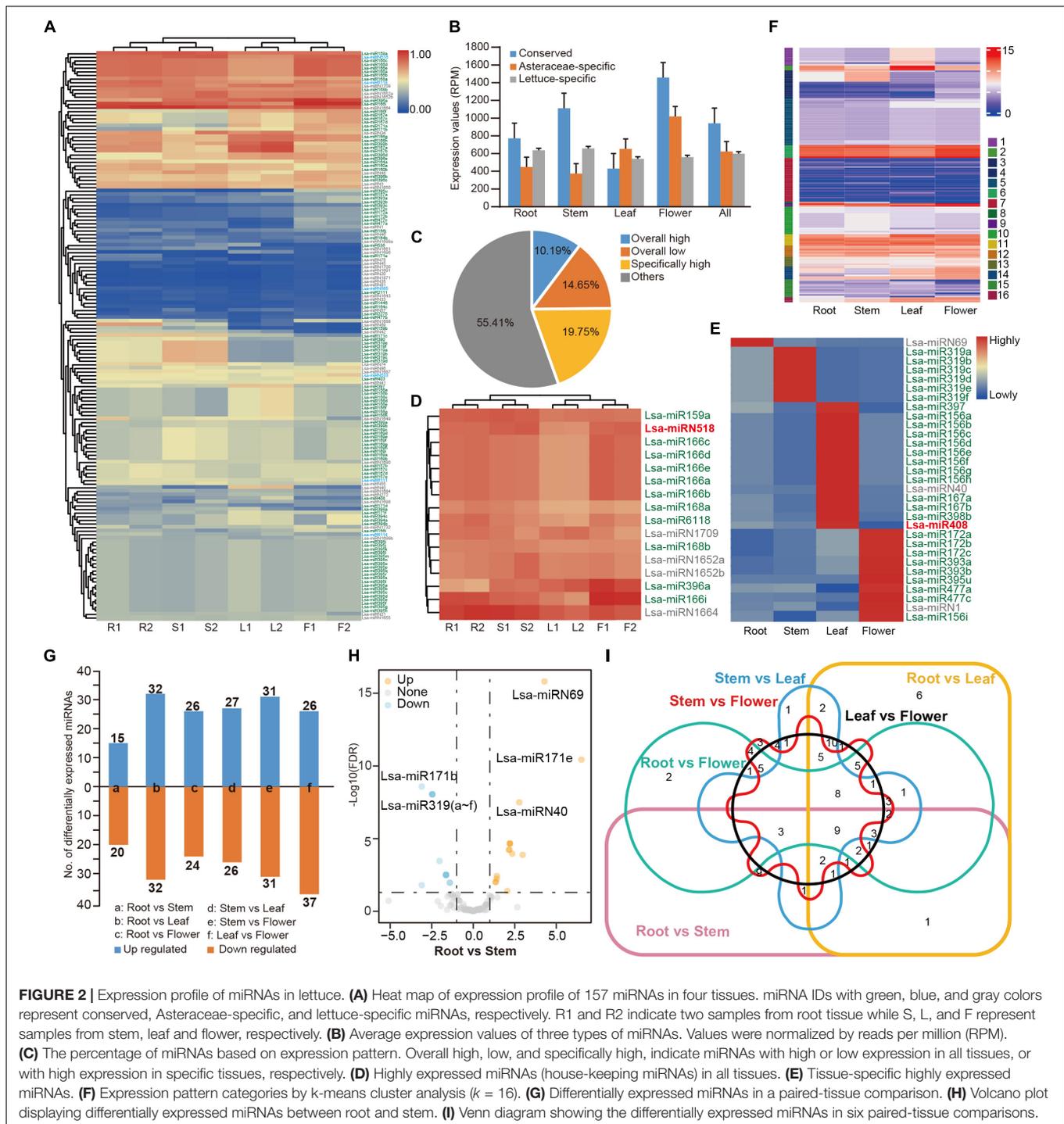
were employed to explore potential miRNA targets. Two computational tools, psRNATarget and RNAhybrid with strict selection criteria (details in Materials and Methods), were first used to predict miRNA targets, with 157 miRNAs and all annotated transcripts of lettuce as inputs. The CleaveLand4.0 (Addo-Quaye et al., 2009) with PARE-Seq data was employed to further detect miRNA targets, where only categories 0–2 were kept for high confidence. As shown in **Figure 3A**, the “union” includes 157 miRNAs and 3,629 transcripts, comprised of 6,540 miRNA-target regulatory pairs when the result from these three methods was united, while the “intersection” possesses 152 miRNAs and 658 transcripts if only the intersected result was kept by at least two methods (**Supplementary Table 10** and **Supplementary Data 1**).

Among the regulatory pairs between these miRNAs and target genes, almost all conserved regulatory relationships that have been studied or reported in model plants, such as Arabidopsis, rice, and maize, were detected, indicating that the predictions are reliable (**Supplementary Table 10**). A number of new regulatory relationships were also discovered. For example, the target genes of conserved miR172 and miR396 can be TFs, ERF and WRKY, respectively, and these predictions are well supported by the PARE-Seq data (**Supplementary Data 1**). In addition, the regulatory relationships between many miRNAs specific to Asteraceae or lettuce and their target genes were discovered. For example, TFs, Dof, and ERF, could potentially be regulated by Lsa-miRN1655 and Lsa-miRN69, which is strongly supported by the evidence (**Supplementary Table 10** and **Supplementary Data 1**). **Figure 3B** shows six examples: miR157a and miR156j are representatives of conserved miRNAs, miRN518 is Asteraceae-specific, while miRN28 is lettuce-specific. All of the regulatory pairs are well supported by PARE-Seq results.

To understand the overall situation of miRNA target genes at the genome-wide level, GO and KEGG enrichment analyses were performed on all miRNA target genes. In GO, biological process (BP) analysis showed the largest enrichment groups were protein binding, protein kinase activity, protein phosphorylation, and signal transduction (**Figure 3C** and **Supplementary Table 11**). In the cellular component analysis, the largest enriched group was nucleus, indicating that a large part of the gene regulated by miRNAs functions in the nucleus. In molecular function analysis, the most abundant genes were related to cellular process, ATP and ADP binding, and various enzyme activities. KEGG enrichment analysis identified signaling pathway or transduction as the main enriched pathways (**Figure 3D** and **Supplementary Table 12**). These are all closely related to miRNAs as regulatory factors and its a large class of target genes, TFs, such as the transmission and transduction of various signals, and functions in the nucleus. The TFs that miRNAs can regulate were investigated further, and 41/23 families with different constrictions were identified (**Figure 3E**).

Expression Correlation Between MicroRNA-Target Regulatory Pairs

Identifying expression patterns of miRNAs and targets in different tissues provides strong evidence for determining



whether there is a real regulatory relationship between them. High-throughput sequencing enables the identification of expression patterns between miRNAs and target genes in batches. To further confirm the regulation between miRNAs and target genes, the mRNAs were sequenced from four tissues, as before. Principal component analysis showed consistency between mRNA and miRNA samples, indicating the reliability of these data (Supplementary Figures 1, 4). The expression patterns

of all potential miRNA targets were obtained from these four tissues. To understand these miRNA-target regulatory pairs, a matrix analysis of the expression patterns of all miRNAs and their targets was pioneered (Figure 4A and Supplementary Table 13). Considering one miRNA can regulate multiple targets, the same target can inversely accept regulations from different miRNAs, and differential spatiotemporal expression exists between miRNAs and targets; thereby a complex matrix

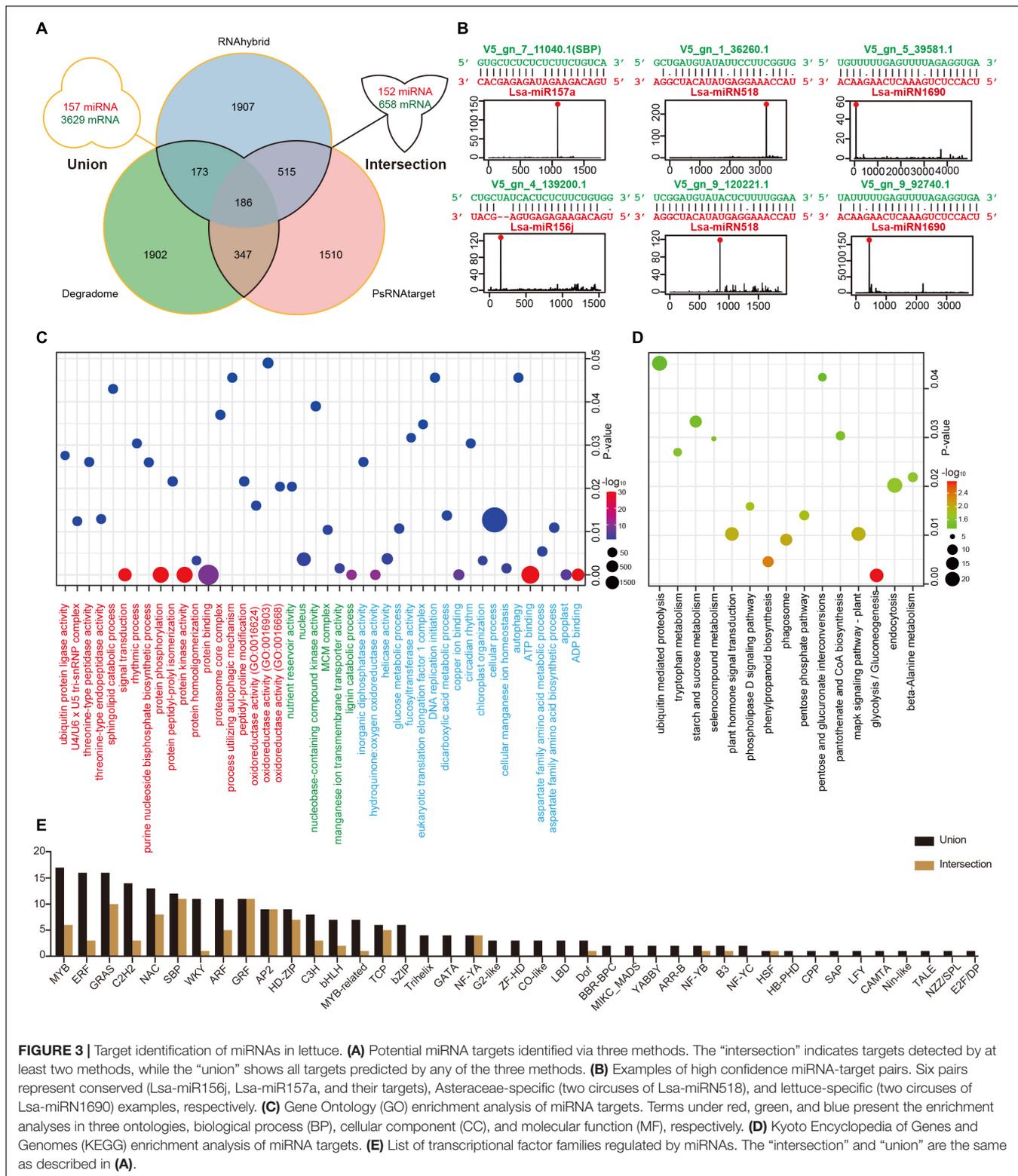
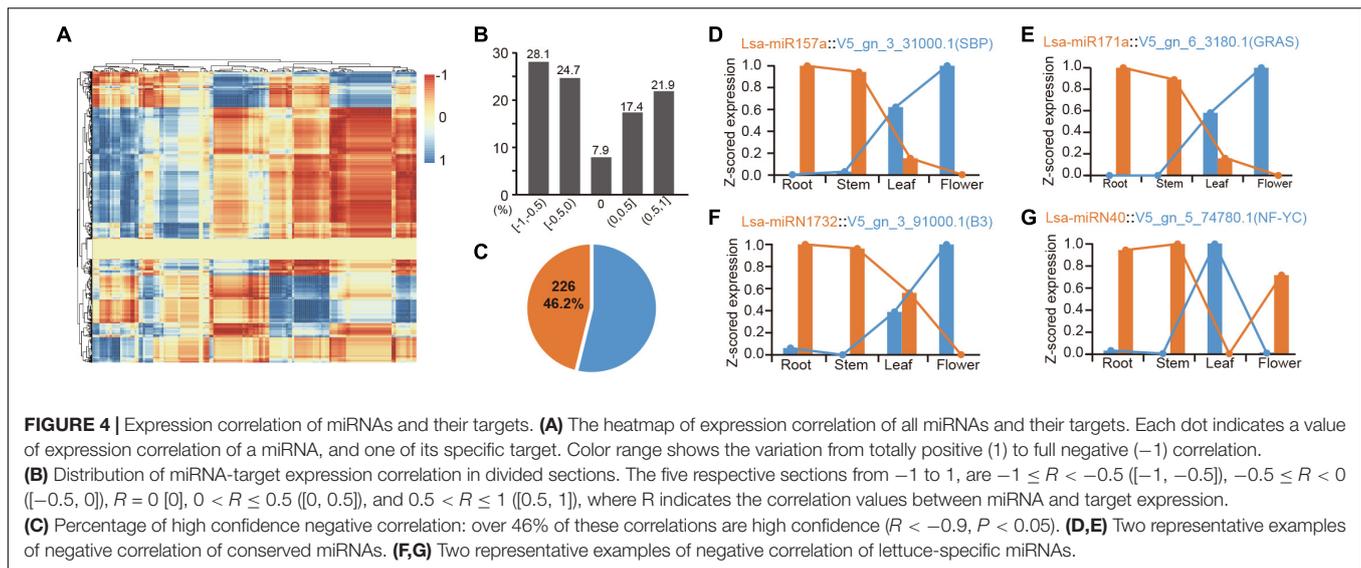


FIGURE 3 | Target identification of miRNAs in lettuce. **(A)** Potential miRNA targets identified via three methods. The “intersection” indicates targets detected by at least two methods, while the “union” shows all targets predicted by any of the three methods. **(B)** Examples of high confidence miRNA-target pairs. Six pairs represent conserved (Lsa-miR156j, Lsa-miR157a, and their targets), Asteraceae-specific (two circuses of Lsa-miRN518), and lettuce-specific (two circuses of Lsa-miRN1690) examples, respectively. **(C)** Gene Ontology (GO) enrichment analysis of miRNA targets. Terms under red, green, and blue present the enrichment analyses in three ontologies, biological process (BP), cellular component (CC), and molecular function (MF), respectively. **(D)** Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis of miRNA targets. **(E)** List of transcriptional factor families regulated by miRNAs. The “intersection” and “union” are the same as described in **(A)**.

was formed (Figure 4A). Interestingly, although these miRNAs and their targets potentially only have a negative regulatory relationship in a specific tissue, the entire matrix showed good support for the negative regulation of miRNAs and target

genes (Figures 4A,B). Data analysis based on “intersection” datasets showed that >46% of the miRNA-target pairs support a strict negative regulatory relationship ($R < 0.9$, $P < 0.05$, Figure 4C).



There are numerous classic cases, such as the expression pattern of miR156 and miR172 in these tissues, which are well supported for their role in developmental transition (Wang et al., 2009; Wu et al., 2009; **Supplementary Table 13**). MiR157a is a good example, where *Lsa-miR157a* and its target *V5_gn_3_31000.1*, a gene harboring SBP (Squamosa-promoter binding protein) domain, have a reverse expression pattern in the four tissues (**Figure 4D**). *Lsa-miR171a* is another representative conserved miRNA, whose target, *V5_gn_6_3180.1*, a GRAS transcriptional factor, has opposite expression patterns (**Figure 4E**). A number of novel regulatory relationships were also revealed, including the relationship between the Asteraceae- and lettuce-specific miRNAs (**Supplementary Table 13**). **Figures 4F,G** lists the two representative lettuce-specific examples that have a strong negative correlation.

Regulatory Network Exploration of Lettuce MicroRNAs

The regulatory elements of miRNAs and target genes do not exist independently, and further constitute a regulatory network. Taking into account the particularity of TFs, that is, they are important miRNA targets that can, in return, regulate miRNA expression, a regulatory network was explored, with TFs, miRNAs, and targets as basic elements. First, miRNA targets from 23 TF families were collected (**Figure 3E**), then the promoter sequences of miRNAs and target genes were extracted, and by a search of known binding motifs of these 23 TFs, a preliminary regulatory relationship of TFs on miRNA and miRNA target genes was constructed. At the same time, using the predicted regulatory relationship between miRNAs and targets, the regulatory relationship between miRNA and target genes was determined (**Figure 5A**). In detail, 1,051 interactions were found between TFs and miRNAs, named TMIs, and 73,133 interactions identified between TFs and miRNA targets, so called TTIs. Together with the 6,540 interactions predicted between miRNAs and targets (MTIs), 46,083 cascades binding TMIs and

MTIs were achieved, and 43,554 FFL motifs were constructed (**Figure 5A**). Based on these FFLs and the links crossing them, a complex regulatory network with TFs, miRNAs, and targets was constructed (**Figure 5B**).

The regulatory network was constructed to further understand the regulation between miRNAs, as the core, and a combination of TFs and target genes. In this network, any element centered on a specific miRNA can independently become a module. For example, miR408 accepts regulations from four TFs, and a dozen targets could be regulated by both miR408 and these four TFs (**Figure 5C**). As described previously, *Lsa-miRN518* is an Asteraceae-specific miRNA (**Figure 2D**), and its 30 targets can also be regulated by several TFs (**Figure 5D**). **Figure 5E** shows a lettuce-specific miRNA, *Lsa-miRN46*, which has a relatively simple FFL module. When the function of each gene was added into these modules, their functions could be further determined. Using **Figure 5F** as an example where all non-TFs and non-miRNA nodes were filtered out, the regulatory network uncovered numerous new links between miRNAs and TFs. In this two particles of the large network, regulations between MYB to miR156, miR156 to SBP, and miR396 to GRF and WRKY, and miR166 to HD-ZIP, have been well established by previous studies (Wu et al., 2009; Zhu et al., 2011; Liebsch and Palatnik, 2020). However, this data supports that more TFs could regulate miR156, and the targets of miR156 could further regulate the modules of miR399 and miRN565. Moreover, more modules are involved in the regulation network of miR396 and miR166, forming a seven-layer network system. Based on the known functions of these miRNAs and TFs, these two network particles, could potentially be involved in development, circadian clock, flowering, response to insect damage and environmental challenges, etc.

Functional Study of Lsa-MIR408

The above expression pattern analysis indicated that *Lsa-miR408* is highly expressed in leaf, while *Lsa-miR408* as a hub gene is in

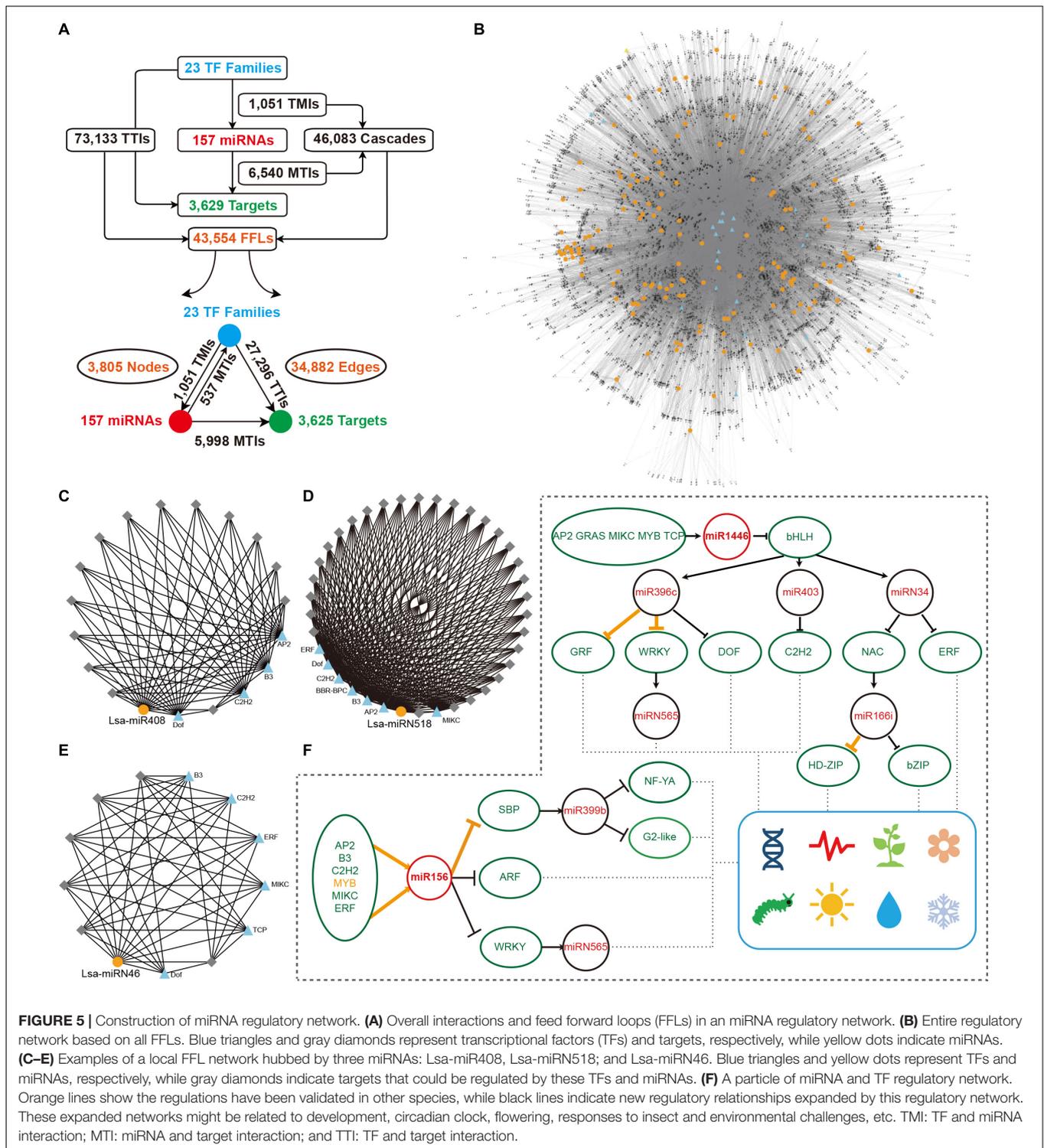


FIGURE 5 | Construction of miRNA regulatory network. **(A)** Overall interactions and feed forward loops (FFLs) in a miRNA regulatory network. **(B)** Entire regulatory network based on all FFLs. Blue triangles and gray diamonds represent transcriptional factors (TFs) and targets, respectively, while yellow dots indicate miRNAs. **(C–E)** Examples of a local FFL network hubbed by three miRNAs: Lsa-miR408, Lsa-miRN518; and Lsa-miRN46. Blue triangles and yellow dots represent TFs and miRNAs, respectively, while gray diamonds indicate targets that could be regulated by these TFs and miRNAs. **(F)** A particle of miRNA and TF regulatory network. Orange lines show the regulations have been validated in other species, while black lines indicate new regulatory relationships expanded by this regulatory network. These expanded networks might be related to development, circadian clock, flowering, responses to insect and environmental challenges, etc. TMI: TF and miRNA interaction; MTI: miRNA and target interaction; and TTI: TF and target interaction.

the miRNA regulatory network (Figure 5E and Supplementary Table 20). In addition, previous (Zhao et al., 2016; Zhang J. P. et al., 2017; Pan et al., 2018) research shows that miR408 is highly conserved in the plant kingdom, and has multiple roles in different life activities, especially in promoting growth which is critical to leafy vegetables. Thus, to further verify the

function of the miRNA regulatory network, miR408 was chosen as an example, and a series of experiments were carried out. Supplementary Table 4 shows that only one miR408 locus was identified, namely Lsa-MIR408, which has a standard stem-loop and a comparable secondary structure to model species, such as *Arabidopsis thaliana* (Supplementary Figure 6A).

Target gene prediction, together with the PARE-Seq experiment, identified that the top three targets of *Lsa-miR408* are copper related genes: basic blue protein (*LsaBBP*), blue copper protein (*LsaBCP*), and copper-transporting ATPase PAA2 (*LsaPAA2*; **Figures 6A–C**, **Supplementary Figure 6B**, and **Supplementary Table 21**). Furthermore, two of the top three targets of *Ath-miR408* are shared with *Arabidopsis* (**Supplementary Figures 6C–E**), suggesting that *Lsa-MIR408* might have similar functions as its companion in *Arabidopsis*. To further explore its function, an *Lsa-miR408* overexpression construct was conducted and introduced into lettuce by *A. tumefaciens*-mediated plant transformation (see section “Materials and Methods”). Compared with wild type plants (WTs), eight independent transgenic lines showed higher accumulation of mature *miR408*, and were designated as OE/408OE (**Figure 6D**). Based on the expression level of *miR408*, three transgenic lines, OE-2, OE-5, and OE-6, were used in subsequent experiments. To test whether the higher expression level of *miR408* is functional, the expression level of three target genes was examined. Quantitative reverse transcription PCR showed that expression of three targets was significantly suppressed in the three independent transgenic lines (**Figure 6E**), indicating that constitutive expression of *miR408* is functional.

Compared to WTs, 408OE plants showed increased growth vigor, exhibiting larger leaves and longer petiole at the seedling stage (**Figures 6F–J**). The petiole of left side blade of 15-day seedlings was used as representatives to quantitatively detect differences in petiole length. Petiole length of 408OE plants increased by 20% (**Figures 6G,H**). At the harvest stage, transgenic plants showed a significant improvement in vegetative development and size (**Figure 6I**). Thus, 50-day 408OE plants were weighed, and were significantly heavier (a 15% increase) than the respective control (**Figure 6J**). Achenes in the transgenic line were also significantly larger than those of WTs (**Figures 6K,L**), and the achenes of 408OE plants were 20% longer than the controls (**Figure 6L**). All of these phenotypic changes are comparable to these in *Arabidopsis* (Jiang et al., 2021) and rice (Zhang J. P. et al., 2017), suggesting the similar molecular mechanism of *miR408* exists in lettuce.

DISCUSSION

The systematic identification and annotation of miRNAs is the first step in studying their functions in a specific species. In this study, based on the lettuce reference genome (Reyes-Chin-Wo et al., 2017), miRNAs in lettuce were systematically annotated by sRNA sequencing and the plant miRNA identification pipeline (Kuang et al., 2019) developed by our laboratory. For the 157 high confidence miRNAs, detailed basic information is provided, such as sequences, locations, structures, miRNA clusters, etc., as well as an analysis of the overall characteristics of these miRNAs, such as the length distribution of precursors, the base preference of the first nt of mature miRNA, conservation between species, and the first attempt to scan their selection after the WGT event. Furthermore, the expression pattern of these miRNAs in four selected tissues was established through sRNA sequencing. The

combination of these analyses provides a strong foundation for understanding and mining miRNA functions in lettuce.

The function of a miRNA is mainly reflected in its targets, and accurate prediction of miRNA targets is critical for exploring miRNA functions. Through the psRNATarget (Dai et al., 2018) and RNAhybrid (Kruger and Rehmsmeier, 2006) pipelines, and high-throughput experimental verification, PARE-Seq, all potential target genes of miRNAs in lettuce were systematically identified. GO and KEGG analyses identified the diversity of these miRNA target genes, and also explained the significance of miRNA functions, being involved in almost all biological activities (Mi et al., 2019). According to the analysis of TFs, many conservative regulations were also detected in lettuce, such as *miR156-SPB*, *miR172-AP2*, etc., confirming reliability of these data. In addition, the expression patterns between miRNAs and their targets in different tissues were correlated: although the expression correlation among them is highly complicated, a general negative expression correlation was observed between miRNAs and their target genes.

In the past few years, plant miRNA research has continued to expand, and new functions of well-studied miRNAs are being continuously discovered. For example, *miR156* research has moved beyond its role in development transition, and new functions in seed dormancy have been discovered (Miao et al., 2019), while the *miR168-AOG1* module has a newly discovered role in plant immunity (Wang H. et al., 2021). At the same time, the study of the miRNA regulation network is becoming increasingly popular because it can promote the discovery of new miRNA functions or regulatory mechanisms. To this end, the miRNA regulatory network in lettuce was constructed, which centered on miRNAs and combined upstream TFs and downstream target genes. Since TFs are also an important class of miRNA target genes, the FFL motifs composed of miRNAs, miRNA targets, and TFs are powerful way to understand the miRNA regulatory network. The regulatory network constructed in this study, not only verified the conserved regulatory relationships found in other species, but discovered numerous novel regulatory modules and potential new connections between different modules. For example, a particle of this network displayed in **Figure 5F** could be involved in many biological activities, as the regulatory network centered on *miR156*, *miR396*, and *miR166* has been expanded.

To further confirm the function of this regulation network, *miR408* was selected as an example, given its role in promoting vegetative growth in model species (Zhao et al., 2016; Zhang J. P. et al., 2017; Pan et al., 2018). As one of the most commercially important leafy vegetables, accelerating vegetative growth and increasing the fresh weight are of great significance to the lettuce industry. Thus, via genetic experiments, the function of *miR408* was verified. Compared with WT lettuce, the fresh weight of the over-expressed *Lsa-MIR408* lines was significantly increased. More interestingly, the plants overexpressing *Lsa-MIR408* can bear larger achenes, which has important potential for lettuce breeding and industrial application. Lettuce achenes in general are very small; therefore, increasing achene size is beneficial to mechanized achene screening and planting, as well as potentially increasing the yield of oil-rich achene

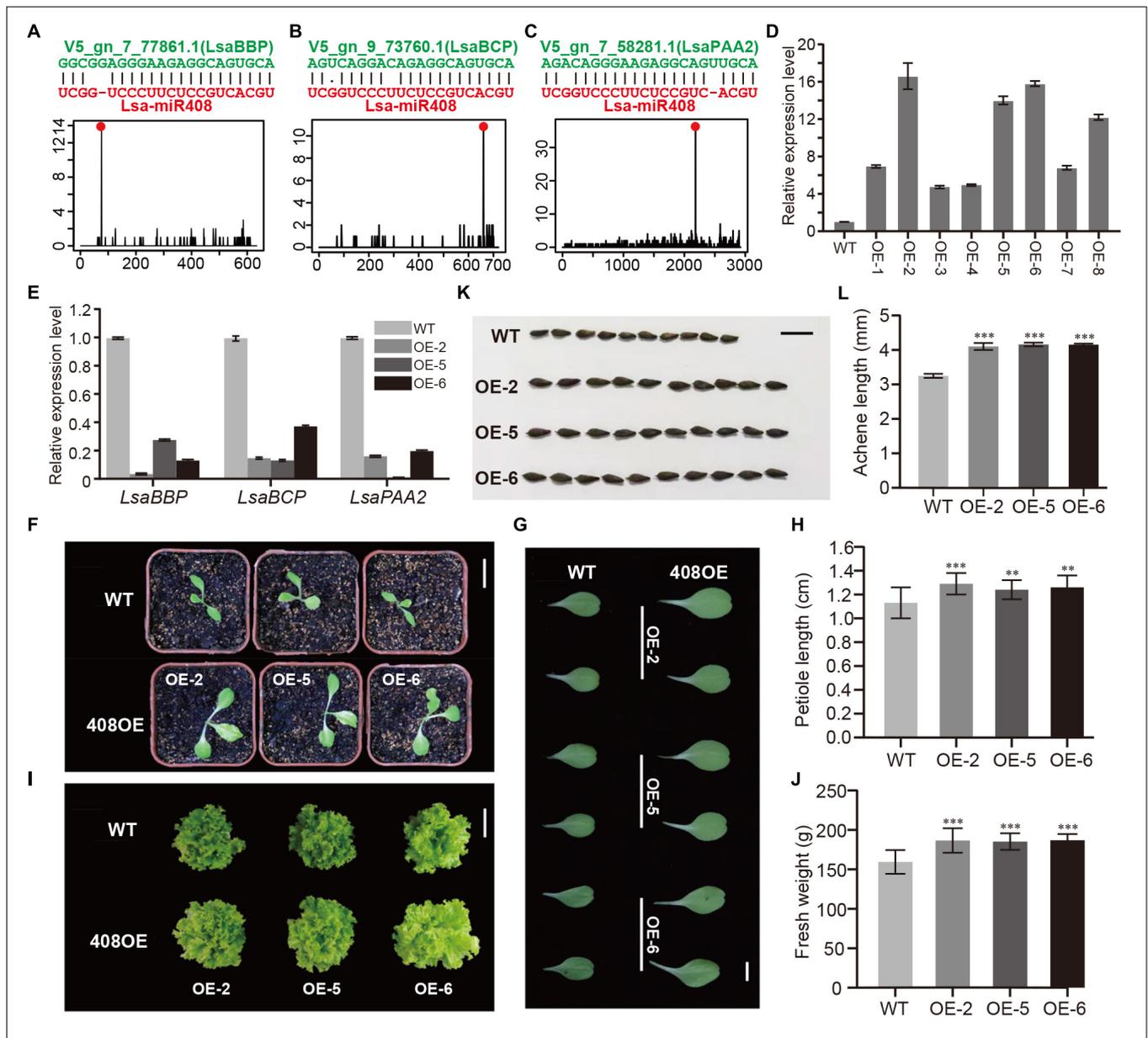


FIGURE 6 | Functional study of Lsa-miR408. **(A–C)** Sequence alignments and degradome profiles of three target genes of Lsa-miR408. The top of each panel shows miRNA-target alignment, while the bottom diagram represents the reads count of degradome data across the target transcript, where the red dot indicates miRNA cleavage site. **(D)** Relative expression levels of miR408 in transgenic lines and wild-type (WT) plants through qRT-PCR analysis. **(E)** Relative expression levels of three target transcripts in three transgenic lines and WT plants by qRT-PCR analysis. **(F)** 12-day seedlings of WT and three transgenic lines, bar = 5 cm. **(G)** Comparison of 12-day cotyledons between WT and three transgenic lines, bar = 1 cm. **(H)** Quantitative measurement of petiole length of 12-day WT and 408OE plants, bar = 1 cm. **(I)** Size comparison of 50-day WT and 408OE plants (top view), bar = 15 cm. **(J)** Quantitative measurement of fresh weight of 50-day WT and miR408-OE plants. **(K)** Comparison of achene morphology and achene size between WT and 408OE plants, bar = 50 mm. **(L)** Quantitative measurement of achene length of WT and 408OE lines. Among these panels, OE/408OE indicates the transgenic lines in which *Lsa-MIR408* was overexpressed. Error bars represent standard deviations (SDs), the number of plants and achenes for quantitative measurement are 20 and 30, respectively, $P < 0.01$ (**), and $P < 0.001$ (***) by Student's *t*-test.

lettuce, as described in watermelon (Wang Y. et al., 2021). The example of miR408 demonstrates the significance of constructing the miRNA regulatory network for understanding miRNA functions.

In brief, as an important economic crop and a representative Asteraceae species, research on the variation in morphology,

color selection, and domesticated trait evolution of lettuce is very limited. Here, miRNAs in lettuce were systematically identified and annotated, and a comprehensive miRNA regulatory network was constructed. This research lays the foundation and provides resources for studying the functions of miRNAs in lettuce, and also provides new ideas for lettuce research and breeding.

DATA AVAILABILITY STATEMENT

The sequencing data for this study can be found in the NCBI Sequence Read Archive (SRA) under accession number PRJNA748591. All annotated miRNA loci were deposited into PmiREN database (<https://www.pmiren.com/>).

AUTHOR CONTRIBUTIONS

XY, LL, and XG initiated and designed the research. YD, ZK, and JD analyzed the data. YQ, PY, YZ, and YW performed the experiments. XY, YD, and PY wrote the manuscript. XY, LL, DL, and JW revised the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Natural Science Foundation of China (NSFC; 32070248 to XY), the Beijing Academy of Agriculture and Forestry Sciences (BAAFS; KJCX20200204 to XY, KJCX201907-2 to JW, KJCX201917 to XG, and QNJJ202019 to YZ), and Peking University Qidong Innovation Project (PUDIP to LL).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.781836/full#supplementary-material>

Supplementary Table 1 | Reads summary in sRNA libraries.

Supplementary Table 2 | Length distribution of clean reads in sRNA libraries.

Supplementary Table 3 | Length distribution of unique reads in sRNA libraries.

REFERENCES

- Addo-Quaye, C., Miller, W., and Axtell, M. J. (2009). CleaveLand: a pipeline for using degradome data to find cleaved small RNA targets. *Bioinformatics* 25, 130–131. doi: 10.1093/bioinformatics/btn604
- Alves, L. Jr., Niemeier, S., Hauenschild, A., Rehmsmeier, M., and Merkle, T. (2009). Comprehensive prediction of novel microRNA targets in *Arabidopsis thaliana*. *Nucleic Acids Res.* 37, 4010–4021. doi: 10.1093/nar/gkp272
- Axtell, M. J., and Meyers, B. C. (2018). Revisiting criteria for plant MicroRNA annotation in the era of big data. *Plant Cell* 30, 272–284. doi: 10.1105/tpc.17.00851
- Bartel, D. P. (2009). MicroRNAs: target recognition and regulatory functions. *Cell* 136, 215–233. doi: 10.1016/j.cell.2009.01.002
- Bartlett, A., O'Malley, R. C., Huang, S. C., Galli, M., Nery, J. R., Gallavotti, A., et al. (2017). Mapping genome-wide transcription-factor binding sites using DAP-seq. *Nat. Protoc.* 12, 1659–1672. doi: 10.1038/nprot.2017.055
- Brown, J., Pirrung, M., and McCue, L. A. (2017). FQC Dashboard: integrates FastQC results into a web-based, interactive, and extensible FASTQ quality control tool. *Bioinformatics* 33, 3137–3139. doi: 10.1093/bioinformatics/btx373
- Carthew, R. W., and Sontheimer, E. J. (2009). Origins and Mechanisms of miRNAs and siRNAs. *Cell* 136, 642–655. doi: 10.1016/j.cell.2009.01.035

Supplementary Table 4 | Summary of 157 miRNAs in lettuce.

Supplementary Table 5 | List of miRNA clusters within 10 kb.

Supplementary Table 6 | List of single miRNAs in synteny blocks.

Supplementary Table 7 | List of paired miRNAs in synteny blocks.

Supplementary Table 8 | Expression profile of 157 miRNAs in lettuce.

Supplementary Table 9 | Differentially expressed miRNAs among tissues.

Supplementary Table 10 | Detailed annotation information and expression profile of 6,540 miRNA-target pairs.

Supplementary Table 11 | The result of gene ontology (GO) enrichment analysis of 3,629 candidate targets.

Supplementary Table 12 | The result of Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis of 3,629 candidate targets.

Supplementary Table 13 | Correlation analysis of expression profile between 157 miRNAs and 3,629 targets.

Supplementary Table 14 | Detailed information of 46,083 Cascade motifs based on “union” datasets.

Supplementary Table 15 | Detailed information of 8,859 Cascade motifs based on “intersection” datasets.

Supplementary Table 16 | Detailed information of 43,554 FFLs based on “union” datasets.

Supplementary Table 17 | Detailed information of 8,424 FFLs based on “intersection” datasets.

Supplementary Table 18 | Regulatory network statistics of 157 miRNAs as hub nodes based on “union” datasets.

Supplementary Table 19 | Regulatory network statistics of 152 miRNAs as hub nodes based on “intersection” datasets.

Supplementary Table 20 | The information between Lsa-miR408 and its target.

Supplementary Table 21 | The result of gene ontology (GO) enrichment analysis of Lsa-miR408 targets.

Supplementary Table 22 | Primers used in this study.

Supplementary Data 1 | Results of categories 0–2 from PARE-Seq analysis (including three subfiles: 1_1, 1_2, 1_3).

- Chandra, S., Vimal, D., Sharma, D., Rai, V., Gupta, S. C., and Chowdhuri, D. K. (2017). Role of miRNAs in development and disease: lessons learnt from small organisms. *Life Sci.* 185, 8–14. doi: 10.1016/j.lfs.2017.07.017
- Chen, Z., Han, Y., Ning, K., Ding, Y., Zhao, W., Yan, S., et al. (2017). Inflorescence development and the role of LsFT in regulating bolting in lettuce (*Lactuca sativa* L.). *Front. Plant Sci.* 8:2248. doi: 10.3389/fpls.2017.02248
- Chen, Z., Zhao, W., Ge, D., Han, Y., Ning, K., Luo, C., et al. (2018). LCM-seq reveals the crucial role of LsSOC1 in heat-promoted bolting of lettuce (*Lactuca sativa* L.). *Plant J* 95, 516–528. doi: 10.1111/tpj.13968
- Cheng, Z., Hou, D., Ge, W., Li, X., Xie, L., Zheng, H., et al. (2020). Integrated mRNA, MicroRNA transcriptome and degradome analyses provide insights into stamen development in Moso Bamboo. *Plant Cell Physiol.* 61, 76–87. doi: 10.1093/pcp/pcz179
- Dai, X., Zhuang, Z., and Zhao, P. X. (2018). psRNATarget: a plant small RNA target analysis server (2017 release). *Nucleic Acids Res.* 46, W49–W54. doi: 10.1093/nar/gky316
- Daniell, H., Kulis, M., and Herzog, R. W. (2019). Plant cell-made protein antigens for induction of Oral tolerance. *Biotechnol. Adv.* 37:107413. doi: 10.1016/j.biotechadv.2019.06.012
- Fromm, B., Keller, A., Yang, X., Friedlander, M. R., Peterson, K. J., and Griffiths-Jones, S. (2020). Quo vadis microRNAs? *Trends Genet* 36, 461–463. doi: 10.1016/j.tig.2020.03.007

- Fujii, H., Chiou, T. J., Lin, S. I., Aung, K., and Zhu, J. K. (2005). A miRNA involved in phosphate-starvation response in Arabidopsis. *Curr. Biol.* 15, 2038–2043. doi: 10.1016/j.cub.2005.10.016
- Gao, Z., Li, J., Li, L., Yang, Y., Li, J., Fu, C., et al. (2021). Structural and Functional Analyses of Hub MicroRNAs in an integrated gene regulatory network of Arabidopsis. *Genomics Proteomics Bioinformatics*. [Epub ahead of print]. doi: 10.1016/j.gpb.2020.02.004
- German, M. A., Luo, S., Schroth, G., Meyers, B. C., and Green, P. J. (2009). Construction of Parallel Analysis of RNA Ends (PARE) libraries for the study of cleaved miRNA targets and the RNA degradome. *Nat. Protoc.* 4, 356–362. doi: 10.1038/nprot.2009.8
- Guo, Z., Kuang, Z., Wang, Y., Zhao, Y., Tao, Y., Cheng, C., et al. (2020). PmiREN: a comprehensive encyclopedia of plant miRNAs. *Nucleic Acids Res.* 48, D1114–D1121. doi: 10.1093/nar/gkz894
- Huo, H., Wei, S., and Bradford, K. J. (2016). DELAY OF GERMINATION1 (DOG1) regulates both seed dormancy and flowering time through microRNA pathways. *Proc. Natl. Acad. Sci. U.S.A.* 113, E2199–E2206. doi: 10.1073/pnas.1600558113
- Inderbitzin, P., Christopoulou, M., Lavelle, D., Reyes-Chin-Wo, S., Michelmore, R. W., Subbarao, K. V., et al. (2019). The LsVeIL allele provides a molecular marker for resistance to *Verticillium dahliae* race 1 in lettuce. *BMC Plant Biol.* 19:305. doi: 10.1186/s12870-019-1905-9
- Jiang, A., Guo, Z., Pan, J., Yang, Y., Zhuang, Y., Zuo, D., et al. (2021). The PIF1-miR408-PLANTACYANIN repression cascade regulates light-dependent seed germination. *Plant Cell* 33, 1506–1529. doi: 10.1093/plcell/koab060
- Jin, J., Tian, F., Yang, D. C., Meng, Y. Q., Kong, L., Luo, J., et al. (2017). PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Res.* 45, D1040–D1045. doi: 10.1093/nar/gkw982
- Johnson, D. S., Mortazavi, A., Myers, R. M., and Wold, B. (2007). Genome-wide mapping of *in vivo* Protein-DNA Interactions. *Science* 316, 1497–1502. doi: 10.1126/science.1141319
- Kalvari, I., Argasinska, J., Quinones-Olvera, N., Nawrocki, E. P., Rivas, E., Eddy, S. R., et al. (2018). Rfam 13.0: shifting to a genome-centric resource for non-coding RNA families. *Nucleic Acids Res.* 46, D335–D342. doi: 10.1093/nar/gkx1038
- Kechin, A., Boyarskikh, U., Kel, A., and Filipenko, M. (2017). cutPrimers: a new tool for accurate cutting of primers from reads of targeted next generation sequencing. *J. Comput. Biol.* 24, 1138–1143. doi: 10.1089/cmb.2017.0096
- Kim, D., Paggi, J. M., Park, C., Bennett, C., and Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* 37, 907–915. doi: 10.1038/s41587-019-0201-4
- Kruger, J., and Rehmsmeier, M. (2006). RNAhybrid: microRNA target prediction easy, fast and flexible. *Nucleic Acids Res.* 34, W451–W454. doi: 10.1093/nar/gkl243
- Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., et al. (2009). Circos: an information aesthetic for comparative genomics. *Genome Res.* 19, 1639–1645. doi: 10.1101/gr.092759.109
- Kuang, Z., Wang, Y., Li, L., and Yang, X. (2019). miRDeep-P2: accurate and fast analysis of the microRNA transcriptome in plants. *Bioinformatics* 35, 2521–2522. doi: 10.1093/bioinformatics/bty972
- Langmead, B. (2010). Aligning short sequencing reads with Bowtie. *Curr. Protoc. Bioinformatics* Chapter 11:Unit 11.7. doi: 10.1002/0471250953.bi1107s32
- Lee, Y., Kim, M., Han, J., Yeom, K. H., Lee, S., Baek, S. H., et al. (2004). MicroRNA genes are transcribed by RNA polymerase II. *EMBO J.* 23, 4051–4060. doi: 10.1038/sj.emboj.7600385
- Liebsch, D., and Palatnik, J. F. (2020). MicroRNA miR396, GRF transcription factors and GIF co-regulators: a conserved plant growth regulatory module with potential for breeding and biotechnology. *Curr. Opin. Plant Biol.* 53, 31–42. doi: 10.1016/j.pbi.2019.09.008
- Mi, H., Muruganujan, A., Ebert, D., Huang, X., and Thomas, P. D. (2019). PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Res.* 47, D419–D426. doi: 10.1093/nar/gky1038
- Miao, C., Wang, Z., Zhang, L., Yao, J., Hua, K., Liu, X., et al. (2019). The grain yield modulator miR156 regulates seed dormancy through the gibberellin pathway in rice. *Nat. Commun.* 10:3822. doi: 10.1038/s41467-019-11830-5
- OGata, H., Goto, S., Sato, K., Fujibuchi, W., Bono, H., and Kanehisa, M. (1999). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 27, 29–34. doi: 10.1093/nar/27.1.29
- Pan, J., Huang, D., Guo, Z., Kuang, Z., Zhang, H., Xie, X., et al. (2018). Overexpression of microRNA408 enhances photosynthesis, growth, and seed yield in diverse plants. *J. Integr. Plant Biol.* 60, 323–340. doi: 10.1111/jipb.12634
- Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T. C., Mendell, J. T., and Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* 33, 290–295. doi: 10.1038/nbt.3122
- Quinlan, A. R., and Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842. doi: 10.1093/bioinformatics/btq033
- Reyes-Chin-Wo, S., Wang, Z., Yang, X., Kozik, A., Arikiti, S., Song, C., et al. (2017). Genome assembly with *in vitro* proximity ligation data and whole-genome triplication in lettuce. *Nat. Commun.* 8:14953. doi: 10.1038/ncomms14953
- Seki, K., Komatsu, K., Tanaka, K., Hiraga, M., Kajiyama-Kanegae, H., Matsumura, H., et al. (2020). A CIN-like TCP transcription factor (*LsTCP4*) having retrotransposon insertion associates with a shift from Salinas type to Empire type in crisphead lettuce (*Lactuca sativa* L.). *Hortic. Res.* 7:15. doi: 10.1038/s41438-020-0241-4
- Shahid, S., Kim, G., Johnson, N. R., Wafula, E., Wang, F., Coruh, C., et al. (2018). MicroRNAs from the parasitic plant *Cuscuta campestris* target host messenger RNAs. *Nature* 553, 82–85. doi: 10.1038/nature25027
- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. doi: 10.1101/gr.1239303
- Su, W., Tao, R., Liu, W., Yu, C., Yue, Z., He, S., et al. (2020). Characterization of four polymorphic genes controlling red leaf colour in lettuce that have undergone disruptive selection since domestication. *Plant Biotechnol. J.* 18, 479–490. doi: 10.1111/pbi.13213
- Tamura, Y., Mori, T., Nakabayashi, R., Kobayashi, M., Saito, K., Okazaki, S., et al. (2018). Metabolomic evaluation of the quality of leaf lettuce grown in practical plant factory to capture metabolite signature. *Front. Plant Sci.* 9:665. doi: 10.3389/fpls.2018.00665
- Tav, C., Tempel, S., Poligny, L., and Tahiri, F. (2016). miRNAfold: a web server for fast miRNA precursor prediction in genomes. *Nucleic Acids Res.* 44, W181–W184. doi: 10.1093/nar/gkw459
- Tian, F., Yang, D. C., Meng, Y. Q., Jin, J., and Gao, G. (2020). PlantRegMap: charting functional regulatory maps in plants. *Nucleic Acids Res.* 48, D1104–D1113. doi: 10.1093/nar/gkz1020
- Tsikou, D., Yan, Z., Holt, D. B., Abel, N. B., Reid, D. E., Madsen, L. H., et al. (2018). Systemic control of legume susceptibility to rhizobial infection by a mobile microRNA. *Science* 362, 233–236. doi: 10.1126/science.aat6907
- Voinnet, O. (2009). Origin, biogenesis, and activity of plant microRNAs. *Cell* 136, 669–687. doi: 10.1016/j.cell.2009.01.046
- Wang, H., Li, Y., Chern, M., Zhu, Y., Zhang, L. L., Lu, J. H., et al. (2021). Suppression of rice miR168 improves yield, flowering time and immunity. *Nat. Plants* 7, 129–136. doi: 10.1038/s41477-021-00852-x
- Wang, J. W., Czech, B., and Weigel, D. (2009). miR156-regulated SPL transcription factors define an endogenous flowering pathway in Arabidopsis thaliana. *Cell* 138, 738–749. doi: 10.1016/j.cell.2009.06.014
- Wang, Y., Tang, H., DeBarry, J. D., Tan, X., Li, J., Wang, X., et al. (2012). MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* 40:e49. doi: 10.1093/nar/gkr1293
- Wang, Y., Wang, J., Guo, S., Tian, S., Zhang, J., Ren, Y., et al. (2021). CRISPR/Cas9-mediated mutagenesis of CLBG1 decreased seed size and promoted seed germination in watermelon. *Hortic. Res.* 8:70. doi: 10.1038/s41438-021-00506-1
- Wei, T., van Treuren, R., Liu, X., Zhang, Z., Chen, J., Liu, Y., et al. (2021). Whole-genome resequencing of 445 *Lactuca* accessions reveals the domestication history of cultivated lettuce. *Nat. Genet.* 53, 752–760. doi: 10.1038/s41588-021-00831-0

- Wu, G., Park, M. Y., Conway, S. R., Wang, J. W., Weigel, D., and Poethig, R. S. (2009). The sequential action of miR156 and miR172 regulates developmental timing in Arabidopsis. *Cell* 138, 750–759. doi: 10.1016/j.cell.2009.06.031
- Yang, C., Li, D., Mao, D., Liu, X., Ji, C., Li, X., et al. (2013). Overexpression of microRNA319 impacts leaf morphogenesis and leads to enhanced cold tolerance in rice (*Oryza sativa* L.). *Plant Cell Environ.* 36, 2207–2218. doi: 10.1111/pce.12130
- Yang, X., Fishilevich, E., German, M. A., Gandra, P., Narva, K. E. J. G. P., and Bioinformatics. (2021). Elucidation of the microRNA transcriptome in western corn rootworm reveals its dynamic and evolutionary complexity. *Genomics Proteomics Bioinformatics*. [Epub ahead of print]. doi: 10.1016/j.gpb.2019.03.008
- Yu, C., Yan, C., Liu, Y., Liu, Y., Jia, Y., Lavelle, D., et al. (2020). Upregulation of a KN1 homolog by transposon insertion promotes leafy head development in lettuce. *Proc. Natl. Acad. Sci. U.S.A.* 117, 33668–33678. doi: 10.1073/pnas.2019698117
- Zhang, J. P., Yu, Y., Feng, Y. Z., Zhou, Y. F., Zhang, F., Yang, Y. W., et al. (2017). MiR408 regulates grain yield and photosynthesis via a phytoeyanin protein. *Plant Physiol.* 175, 1175–1185. doi: 10.1104/pp.17.01169
- Zhang, L., Su, W., Tao, R., Zhang, W., Chen, J., Wu, P., et al. (2017). RNA sequencing provides insights into the evolution of lettuce and the regulation of flavonoid biosynthesis. *Nat. Commun.* 8:2264. doi: 10.1038/s41467-017-02445-9
- Zhang, S., Sang, X., Hou, D., Chen, J., Gu, H., Zhang, Y., et al. (2019). Plant-derived RNAi therapeutics: a strategic inhibitor of HBsAg. *Biomaterials* 210, 83–93. doi: 10.1016/j.biomaterials.2019.04.033
- Zhang, W., Alseikh, S., Zhu, X., Zhang, Q., Fernie, A. R., Kuang, H., et al. (2020). Dissection of the domestication-shaped genetic architecture of lettuce primary metabolism. *Plant J.* 104, 613–630. doi: 10.1111/tpj.14950
- Zhao, X. Y., Hong, P., Wu, J. Y., Chen, X. B., Ye, X. G., Pan, Y. Y., et al. (2016). The tae-miR408-Mediated Control of TaTOC1 genes transcription is required for the regulation of heading time in wheat. *Plant Physiol.* 170, 1578–1594. doi: 10.1104/pp.15.01216
- Zhao, Y., Kuang, Z., Wang, Y., Li, L., and Yang, X. (2021). MicroRNA annotation in plants: current status and challenges. *Brief. Bioinform.* 22:bbab075. doi: 10.1093/bib/bbab075
- Zhu, H., Hu, F., Wang, R., Zhou, X., Sze, S. H., Liou, L. W., et al. (2011). Arabidopsis Argonaute10 specifically sequesters miR166/165 to regulate shoot apical meristem development. *Cell* 145, 242–256. doi: 10.1016/j.cell.2011.03.024
- Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Deng, Qin, Yang, Du, Kuang, Zhao, Wang, Li, Wei, Guo, Li and Yang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.