BMC Plant Biology

RESEARCH Open Access

# DNA polymorphisms and $N^6$-methyladenosine RNA methylation influencing microRNA expression and regulatory activity in plants

Zedi Feng[1†], Jia Yao[1†], Zhifang Jiang[1], Xiaomei Wu[1], Jiacen Wang[1], Wenyuan Wu[2], Chaogang Shao[3*], Xiaoxia Ma[4*] and Yijun Meng[1*]

## Abstract

Plant subspecies that have colonized distinct natural habitats must evolve different phenotypic and physiological features to ensure survival. However, the underlying molecular mechanisms remain poorly understood. In this study, we selected seven *Arabidopsis* accessions and six rice varieties with significantly different natural habitat parameters to investigate the molecular mechanisms behind the differential expression and activity of microRNAs (miRNAs) at the subspecies level. Compared to Col-0, 24–35 differentially expressed miRNA precursors and 36–110 differentially expressed mature miRNAs were identified in the other six *Arabidopsis* accessions. Compared to Nipponbare, 19–40 differentially expressed precursors and 8–101 mature miRNAs were identified in the other five rice varieties. The expression patterns of the precursors correlate well with those of the mature miRNAs. Both $m^6A$ modification and single nucleotide polymorphisms (SNPs) associated with miRNA genes were closely linked to precursor transcription and miRNA maturation. Degradome sequencing data analysis revealed that the high miRNA level likely caused intense target cleavages in a specific subspecies. In some cases, the affinity of miRNA—target interactions was significantly influenced by SNPs. Notably, many target genes were functionally involved in organ development, reproduction or stress responses. The data presented here provide molecular hints into the different developmental processes and environmental responses among plant subspecies with distinct natural habitats.

**Keywords** MicroRNAs (miRNAs), Single nucleotide polymorphisms (SNPs), $N^6$-methyladenosine ($m^6A$), Differential expression, Targets

[†]Zedi Feng and Jia Yao contributed equally to this work.

*Correspondence:
Chaogang Shao
shaocg@zjhu.edu.cn
Xiaoxia Ma
xiaoma_0910@163.com
Yijun Meng
mengyijun@zju.edu.cn

[1]College of Life and Environmental Sciences, Hangzhou Normal University, Hangzhou 311121, China
[2]School of Information Science and Technology, Hangzhou Normal University, Hangzhou 311121, China
[3]College of Life Sciences, Huzhou Normal University, Huzhou 313000, China
[4]School of Pharmacy, Hangzhou Normal University, Hangzhou 311121, China

## Introduction

As sessile organisms, plants were forced to develop diverse molecular strategies to perceive and respond to the environmental fluctuations. Even for a plant species colonized in distinct natural habitats, several subspecies with different but inheritable phenotypic and physiological features might have evolved for survival, growth, development and reproduction. From another perspective, these subspecies could be the ideal objects for the studies on the influence of the exiguous variations on the elaborate gene networks involved in environmental adaptation, considering their basically homologous genetic backgrounds. Indeed, tremendous efforts have been made to identify DNA polymorphisms, including single nucleotide polymorphisms (SNPs), insertions and deletions, in numerous plant genomes. The first international ambitious attempt for genome-wide discovery of the plant genomic variations was launched at the beginning of 2008. The 1001 Genomes Project (https://1001genomes.org/) investigated more than one thousand accessions of *Arabidopsis* (*Arabidopsis thaliana*) from a worldwide hierarchical collection. By 2016, a total of 1,135 accession genomes along with the sequence polymorphism data had been publicly available [1]. For another model plant rice (*Oryza sativa*), the 3000 Rice Genome Project aimed to re-sequence the genomes of 3,024 rice cultivars collected from 89 Asian countries, and the sequence polymorphism data were made available at https://registry.opendata.aws/3kricegenome/ [2].

MicroRNAs (miRNAs) are one of the well-characterized small RNA species in plants [3]. The mature miRNAs are processed from the stem-loop-structured precursors through dicer-like 1 (DCL1)-mediated two-step cropping. The processing efficiency is influenced by several key factors, such as the components of the miRNA processor and the precursor structures. After maturation, the 21-nt mature miRNA started with 5'U is preferentially incorporated into the Argonaute 1 (AGO1)-centered RNA-induced silencing complex. Then, the mature miRNA guides the silencing complex to the target transcript based on the highly complementary miRNA binding site (MBS). In many cases, the endonucleolytic activity of AGO1 leads to the site-specific target cleavages, which can be detected by modified 5' RACE (rapid amplification of cDNA ends) at fine-scale [4] or degradome sequencing (degradome-seq) at large-scale [5, 6] respectively. The miRNA-guided target cleavages form a post-transcriptional regulatory pathway critical for plant development, hormone signaling and stress responses [7].

Since differential miRNA expression makes a great contribution to the distinct plant characteristics related to development and environmental adaptability, many efforts have been devoted to uncovering the key factors influencing miRNA expression and regulatory activity. Firstly, both natural variations [8] and artificially introduced mutations [9] on the miRNA precursors, especially within the miRNA: miRNA$^*$-coding regions, can perturb the stem-loop structures and affect miRNA maturation. In addition to sequence variations, $N^6$-methyladenosine ($m^6A$) modification, one of the most prevalent RNA methylation types, was also reported to play a critical regulatory role in miRNA biogenesis in both plants and animals [10–13]. Secondly, as mentioned above, the high sequence complementarity between MBSs and mature miRNAs is the prerequisite for efficient target regulation. Several reports have showed the potential impact of the sequence variations of MBSs or mature miRNAs on miRNA—target interactions [14–16]. To date, however, the molecular mechanisms underlying differential miRNA expression and activity, and the links to the differential phenotypic and physiological features are still under investigation in plants.

In the present study, seven *Arabidopsis* accessions (Altai-5, Col-0, ICE73, Kas-2, Kondara, Se-0 and Zal-1) and six rice varieties (93 – 11, DANG YU 5 HAO abbreviated as DY5, MIN BEI WAN XIAN abbreviated as MBWX, NAN TE HAO abbreviated as NTH, Nipponbare abbreviated as NIP, Suweon 295 abbreviated as SW295) were selected for the study on the molecular mechanisms underlying differential miRNA expression and target regulation. Treating Col-0 and NIP as the references in *Arabidopsis* and rice respectively, each of the other subspecies has at least two significantly different natural habitat parameters related to precipitation, temperature or solar radiation. Transcriptome sequencing identified 24, 24, 35, 33, 27 and 32 differentially expressed precursors ($|\log_2 FC| \geq 1$, $FDR < 0.05$) in Altai-5, ICE73, Kas-2, Kondara, Se-0 and Zal-1, as compared to Col-0 respectively, and 20, 40, 32, 27 and 19 differentially expressed precursors ($|\log_2 FC| \geq 1$, $FDR < 0.05$) in 93 – 11, DY5, MBWX, NTH and SW295, as compared to NIP respectively. Small RNA sequencing identified 110, 36, 50, 101, 46 and 50 differentially expressed mature miRNAs ($|\log_2 FC| \geq 1$, $FDR < 0.05$) in Altai-5, ICE73, Kas-2, Kondara, Se-0 and Zal-1, as compared to Col-0 respectively, and 15, 101, 21, 8, 49 differentially expressed mature miRNAs ($|\log_2 FC| \geq 1$, $FDR < 0.05$) in 93 – 11, DY5, MBWX, NTH and SW295, as compared to NIP respectively. In most cases, the differential expression patterns of the precursors positively correlate with those of the mature miRNAs in the subspecies investigated. Further analyses revealed that both $m^6A$ modification and SNPs associated with the miRNA genes were closely linked to precursor transcription and miRNA maturation. Degradome sequencing data analysis revealed that the high mature miRNA level likely caused intense target cleavages in a specific subspecies. Additionally, the

affinity of miRNA—target interactions was significantly influenced by the SNPs resided within the mature miRNAs or the MBSs. In many cases, the target genes were functionally involved in organ development, reproduction or stress responses. Taken together, our results uncovered the effects of both genomic and epitranscriptomic variations on differential miRNA expression and regulation, which provided the molecular insights into the divergent phenotypic and physiological features of the plant subspecies with distinct habitats.

## Materials and methods

### Plant materials and growth conditions

The seeds of seven *Arabidopsis* accessions were surface-sterilized in 70% ethanol for 3 min and then in 10% NaClO for 10 min. It was followed by four washes with sterile distilled water. The seeds were sown on the solid Gamborg B5 medium at pH 5.8 under sterile condition, and incubated at 4℃ in darkness for two days. Then, all of the accessions were moved to a growth chamber at 24℃ for 16 h in the light (150 μmol/$m^2$/s). The 16-day-old seedlings were used for all of the sequencing experiments.

The seeds of six rice varieties were incubated at 38℃ in darkness for five days. Then, the seeds were surface-sterilized in 30% NaClO for 10 min and in 70% ethanol for 1 min. It was followed by five washes with sterile distilled water. The seeds were sown on the solid Murashige and Skoog medium at pH 5.8 under sterile condition, and moved to a growth chamber at 30 °C for 14 h in the light (200 μmol/$m^2$/s). The 10-day-old seedlings were used for all of the sequencing experiments.

### Transcriptome sequencing (RNA-seq) and data analysis

Total RNAs were isolated by using the TRIzol reagent (Invitrogen, USA) following the user's manual. The amount and purity of each RNA sample was quantified by using NanoDrop ND-1000 (NanoDrop, Wilmington, DE, USA). The integrity of each RNA sample was assessed by using Bioanalyzer 2100 (Agilent, CA, USA) under "RIN (RNA integrity number) value > 7.0", and was confirmed by denaturing agarose electrophoresis. The poly(A)-tailed RNAs were purified from 1 μg total RNAs by using Dynabeads Oligo (dT)25-61005 (Thermo Fisher, CA, USA) through two rounds of purification. Then, the poly(A)-tailed RNAs were subject to fragmentation by using Magnesium RNA Fragmentation Module (NEB, cat.e6150, USA) under 94℃ for 7 min. The fragmented RNAs were reverse-transcribed to create the single-stranded cDNAs by using SuperScript™ II Reverse Transcriptase (Invitrogen, cat. 1896649, USA). These single-stranded cDNAs were treated as the templates to synthesize U-labeled double-stranded DNAs by using *E. coli* DNA polymerase I (NEB, cat.m0209, USA), RNase H (NEB, cat.m0297, USA) and dUTP Solution (Thermo Fisher, cat.R0133, USA). Then, an A-base was added to the blunt ends of each strand, preparing them for ligation to the indexed adapters. Size selection was performed for the fragments ligated with the adapters by using AMPure XP beads. After the heat-labile UDG enzyme (NEB, cat. m0280, USA) treatment, the ligated products were amplified by PCR under the following conditions: initial denaturation at 95℃ for 3 min; eight cycles of denaturation at 98℃ for 15 s, annealing at 60℃ for 15 s, and extension at 72℃ for 30 s; and final extension at 72℃ for 5 min. The averaged size of the final cDNA library was 300 ± 50 bp. The paired-end sequencing (PE150) was performed by LC-Bio Technology CO., Ltd. (Hangzhou, China) with the illumina Novaseq™ 6000 platform following the vendor's recommended protocol. For RNA-seq, there are three biological replicates for each subspecies. Each biological replicate is prepared by pooled sampling with more than 100 seedlings.

The pre-treatment of the RNA-seq data, including adapter trimming, low-quality read removal and sequence quality verification, was performed by using fastp with default parameters [17]. HISAT2 [18] was used for mapping the sequencing reads to the miRNA precursors, followed by expression level calculation in FPKM (fragments per kilobase of exon per million fragments mapped). The miRNA precursors differentially expressed in one of the *Arabidopsis* accessions or one of the rice varieties, as compared to Col-0 or NIP respectively, were identified by using the estimateDisp module of edgeR ( http://bioconductor.org/packages/release/bioc/html/edgeR.html) [19]. The glmFit module of edgeR was used for statistical significance test. The cutoff "$|\log_2 FC| \geq 1$, FDR < 0.05" was applied for the differentially expressed precursor screening. The heatmaps were drawn by using the ComplexHeatmap package [20].

### Small RNA and degradome sequencing, and data processing

TruSeq® Small RNA Sample Prep Kit (Illumina, San Diego, USA) was used for small RNA sequencing (sRNA-seq) library construction according to the manufacturer's manual. Briefly, the RNAs less than 200 nt were separated from 1 μg total RNAs by polyacrylamide gel electrophoresis, and were ligated to the adapters by T4 RNA ligase at both ends. By using SuperScript™ II Reverse Transcriptase (Invitrogen, USA), the sRNAs containing the adapters were transcribed to the single-stranded cDNAs. These cDNAs were treated as the PCR templates to generate the double-stranded cDNAs which were then purified by 16% TBE gel electrophoresis. Sequencing was performed by LC-Bio (Hangzhou, China) with the Illumina HiSeq 2500 platform following the vendor's recommended protocol. For sRNA-seq, there are three

biological replicates for each subspecies. Each biological replicate is prepared by pooled sampling with more than 100 seedlings.

For degradome library construction, 150 ng poly(A)-tailed RNAs were obtained from 20 μg total RNAs through two-round purification by using Dynabeads Oligo (dT)25-61005 (Thermo Fisher, CA, USA). Then, the adapters were ligated to the 5' ends of the RNA remnants by RNA ligase. Reverse transcription was performed to create the first strands of cDNAs with a random primer. After size selection with AMPure XP beads, the single-stranded cDNAs were PCR amplified to produce the double-stranded cDNAs. Single-end sequencing (50 bp) was performed by LC-Bio (Hangzhou, China) with the Illumina HiSeq 2500 platform following the vendor's recommended protocol.

For the sRNA- and degradome-seq data, FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/) was used for adapter trimming and low-quality read removal. Then, the pre-treated sRNA sequences ranging from 15 to 45 nt were retained for further analyses, while the pre-treated degradome sequences from 18 to 25 nt were retained. For a sRNA or degradome dataset, the expression level of one sequence was normalized by dividing the raw read count of this sequence by the total raw read counts of this dataset, and then multiplied by $10^6$. The expression levels normalized in RPM (reads per million) enabled cross-library comparisons.

The miRNAs differentially expressed in one of the *Arabidopsis* accessions or one of the rice varieties, as compared to Col-0 or NIP respectively, were identified by using the estimateDisp module of edgeR (http://bioconductor.org/packages/release/bioc/html/edgeR.html) [19]. The glmFit module of edgeR was used for statistical significance test. The cutoff "$|\log_2 FC| \geq 1$, FDR < 0.05" was applied for the differentially expressed miRNA screening. "Omicstudio" (https://www.omicstudio.cn/) [21, 13] was used to draw the Venn diagrams, and the ComplexHeatmap package (https://github.com/jokergoo/ComplexHeatmap) [20] was used to draw the heatmaps.

The miRU algorithm [22, 23] was employed for miRNA target prediction. The predicted miRNA—target pairs were subject to degradome-based validation as described in our previous studies [5, 6]. The secondary structures of the miRNA precursors were predicted by using the RNAfold program of the Vienna RNA package (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi) [24].

#### Methylated RNA immunoprecipitation sequencing (MeRIP-seq) and data analysis

Total RNAs were isolated by using the TRIzol reagent (Invitrogen, USA) following the manufacturer's instructions. The amount and purity of each RNA sample was quantified by using NanoDrop ND-1000 (NanoDrop, USA). The integrity of each RNA sample was assessed by using Bioanalyzer 2100 (Agilent, CA, USA) under the cutoff "RIN (RNA integrity number) value > 7.0", and was confirmed by denaturing agarose electrophoresis. The poly(A)-tailed RNAs were extracted from 30 μg total RNAs by using Dynabeads Oligo (dT)25-61005 (Thermo Fisher, USA) through two rounds of purification. The poly(A)-tailed RNAs were fragmented into small pieces by using Magnesium RNA Fragmentation Module (NEB, cat.e6150, USA) at 86℃ for 7 min. The fragmented RNAs were incubated with the $m^6A$-specific antibody (Synaptic Systems, cat.202003, Germany) in IP buffer (50 mM Tris-HCl, 750 mM NaCl and 0.5% Igepal CA-630) at 4℃ for 2 h. The immunoprecipitated RNAs were reverse-transcribed to create the single-stranded cDNAs by using SuperScript™ II Reverse Transcriptase (Invitrogen, cat.1896649, USA). Next, the single-stranded cDNAs were treated as the templates to synthesize the U-labeled double-stranded DNAs with *E. coli* DNA polymerase I (NEB, cat.m0209, USA), RNase H (NEB, cat.m0297, USA) and dUTP Solution (Thermo Fisher, cat.R0133, USA). An A-base was then added to the blunt ends of each strand, which was ready for ligation to the indexed adapter containing a T-base overhang. After adapter ligation, the DNA fragments were subject to size selection by using the AMPureXP beads. After the heat-labile UDG enzyme (NEB, cat.m0280, USA) treatment, the ligated products were amplified with PCR under the following conditions: initial denaturation at 95℃ for 3 min; 8 cycles of denaturation at 98℃ for 15 s, annealing at 60℃ for 15 s, and extension at 72℃ for 30 s; and final extension at 72℃ for 5 min. The averaged size of the final cDNA library was 300 ± 50 bp. The 2 × 150 bp paired-end sequencing (PE150) was performed by LC-Bio Technology CO., Ltd. (Hangzhou, China) with the illumina Novaseq™ 6000 platform following the vendor's recommended protocol. For $m^6A$ MeRIP-seq, there are three biological replicates for each sample. Each biological replicate is prepared by pooled sampling with more than 100 seedlings.

The pre-treatment of the MeRIP-seq data, including adapter trimming, low-quality read removal and sequence quality verification, was performed by using fastp [17]. HISAT2 [18] was used for mapping the sequencing reads to the reference genomes of *Arabidopsis* (The *Arabidopsis* Information Resource, TAIR10) (https://www.arabidopsis.org/) [25] and rice (the rice genome annotation project, RGAP7.0) (https://rice.uga.edu/) [26] respectively. For both the IP and input libraries, all reads mapped onto the reference genomes were included in $m^6A$ peak calling and differential peak identification by using exomePeak2 (https://rdrr.io/github/ZhenWei10/exomePeak2/) [27, 28]. Considering the non-coding feature of the miRNA genes, the "whole-genome"

mode of exomePeak2 was used. All of the differential peaks were identified under the cutoff "$|\log_2 FC| > 0.5$, $P_{adj} < 0.05$". The resulting peak files in bed format are compatible for visualization on the IGV (Integrative genomics viewer) software (http://www.igv.org) [29]. Besides, based on the "input" datasets, featureCounts [30] was used to calculate the expression levels of the miRNA precursors in FPKM.

## Results

### Differential expression of miRNA precursors and mature miRNAs in distinct subspecies

MiRNAs play an indispensible role in plant development and stress response mostly through post-transcriptional gene regulation [31]. The miRNA expression levels make a great contribution to the distinguishable phenotypic or physiological features, such as organ development, reproduction and environmental adaptability, even at the subspecies level. To investigate the detailed molecular mechanisms, seven *Arabidopsis* accessions (Altai-5, Col-0, ICE73, Kas-2, Kondara, Se-0 and Zal-1) and six rice varieties (93 – 11, DY5, MBWX, NTH, NIP and SW295) were selected for this study. Notably, by treating Col-0 and NIP as the references in *Arabidopsis* and rice respectively, each of the other subspecies was required to have at least two significantly different natural habitat parameters, including annual precipitation, annual mean temperature, annual temperature variation range, minimal solar radiation, maximal solar radiation, and annual mean solar radiation (see details in Fig. S1).

Since the mature miRNAs are processed from the stem-loop precursors [3], the precursor levels have a direct impact on the miRNA abundances. In this regard, RNA-seq with three biological replicates was performed for each subspecies to measure the abundances of the miRNA precursors in FPKM (fragments per kilobase of exon per million fragments mapped). Col-0 and NIP were treated as the control materials for the pairwise comparisons in *Arabidopsis* and rice respectively. According to miRBase (release 22) [32], there are 326 miRNA precursors in *Arabidopsis*. The RNA-seq data analysis revealed 24 (11 up-regulated and 13 down-regulated), 24 (9 up-regulated and 15 down-regulated), 35 (17 up-regulated and 18 down-regulated), 33 (14 up-regulated and 19 down-regulated), 27 (11 up-regulated and 16 down-regulated) and 32 (15 up-regulated and 17 down-regulated) differentially expressed precursors in Altai-5, ICE73, Kas-2, Kondara, Se-0 and Zal-1, respectively ($|\log_2 FC| \geq 1$, FDR < 0.05; Fig. 1A and Table S1). Compared to Col-0, seven precursors are differentially expressed in all of the six accessions investigated. Notably, five out of the seven precursors, i.e. ath-MIR413, ath-MIR5014b, ath-MIR5014a, ath-MIR3932b and ath-MIR841b, are uniformly down-regulated, and one precursor ath-MIR396b are uniformly up-regulated in the six accessions (Table S1). Besides, among the differentially expressed precursors, ath-MIR853 was specifically identified in Altai-5, ath-MIR169m, ath-MIR157b, ath-MIR775 and ath-MIR5023 were specifically identified in ICE73, ath-MIR156d, ath-MIR864, ath-MIR5629 and ath-MIR319a were specifically identified in Kas-2, ath-MIR160b, ath-MIR156j, ath-MIR156a, ath-MIR5630b and ath-MIR5655 were specifically identified in Kondara, ath-MIR4221, ath-MIR822, ath-MIR5020b, ath-MIR5029, ath-MIR8176 and ath-MIR824 were specifically identified in Se-0, and ath-MIR163, ath-MIR5642a, ath-MIR397b and ath-MIR398a were specifically identified in Zal-1 (Fig. 1A and Table S1). In rice, 604 miRNA precursors were registered in miRBase (release 22). Our analysis uncovered 20 (9 up-regulated and 11 down-regulated), 40 (24 up-regulated and 16 down-regulated), 32 (20 up-regulated and 12 down-regulated), 27 (9 up-regulated and 18 down-regulated) and 19 (10 up-regulated and 9 down-regulated) differentially expressed precursors in 93 – 11, DY5, MBWX, NTH and SW295, respectively ($|\log_2 FC| \geq 1$, FDR < 0.05; Fig. 1C and Table S2). Compared to NIP, nine precursors are differentially expressed in all of the five varieties investigated. Notably, four out of the nine precursors, i.e. osa-MIR397b, osa-MIR398b, osa-MIR408 and osa-MIR528, are uniformly up-regulated, and the remaining five precursors, i.e. osa-MIR815b, osa-MIR2103, osa-MIR169b, osa-MIR7695 and osa-MIR6247, are uniformly down-regulated in the five varieties (Table S2). Besides, among the differentially expressed precursors, osa-MIR169c was specifically discovered in 93 – 11, osa-MIR398a, osa-MIR171h, osa-MIR171f, osa-MIR164a, osa-MIR2863b, osa-MIR160b, osa-MIR156c, osa-MIR5493, osa-MIR2055, osa-MIR156b, osa-MIR5801b, osa-MIR160a, osa-MIR5531 and osa-MIR1320 were specifically discovered in DY5, osa-MIR7693, osa-MIR1431, osa-MIR6256, osa-MIR5504, osa-MIR162a and osa-MIR167h were specifically discovered in MBWX, osa-MIR6253, osa-MIR812j, osa-MIR6255, osa-MIR166a, osa-MIR1868, osa-MIR6246, osa-MIR5535 and osa-MIR171c were specifically discovered in NTH, and osa-MIR418 was specifically discovered in SW295 (Fig. 1C and Table S2).

To calculate the mature miRNA abundances in RPM (reads per million), sRNA-seq with three biological replicates was performed for each subspecies. According to miRBase (release 22), there are 428 mature miRNAs in *Arabidopsis*. Compared to Col-0, the sRNA-seq data analysis revealed 110 (33 up-regulated and 77 down-regulated), 36 (16 up-regulated and 20 down-regulated), 50 (9 up-regulated and 41 down-regulated), 101 (26 up-regulated and 75 down-regulated), 46 (12 up-regulated and 34 down-regulated) and 50 (20 up-regulated and 30 down-regulated) differentially expressed mature

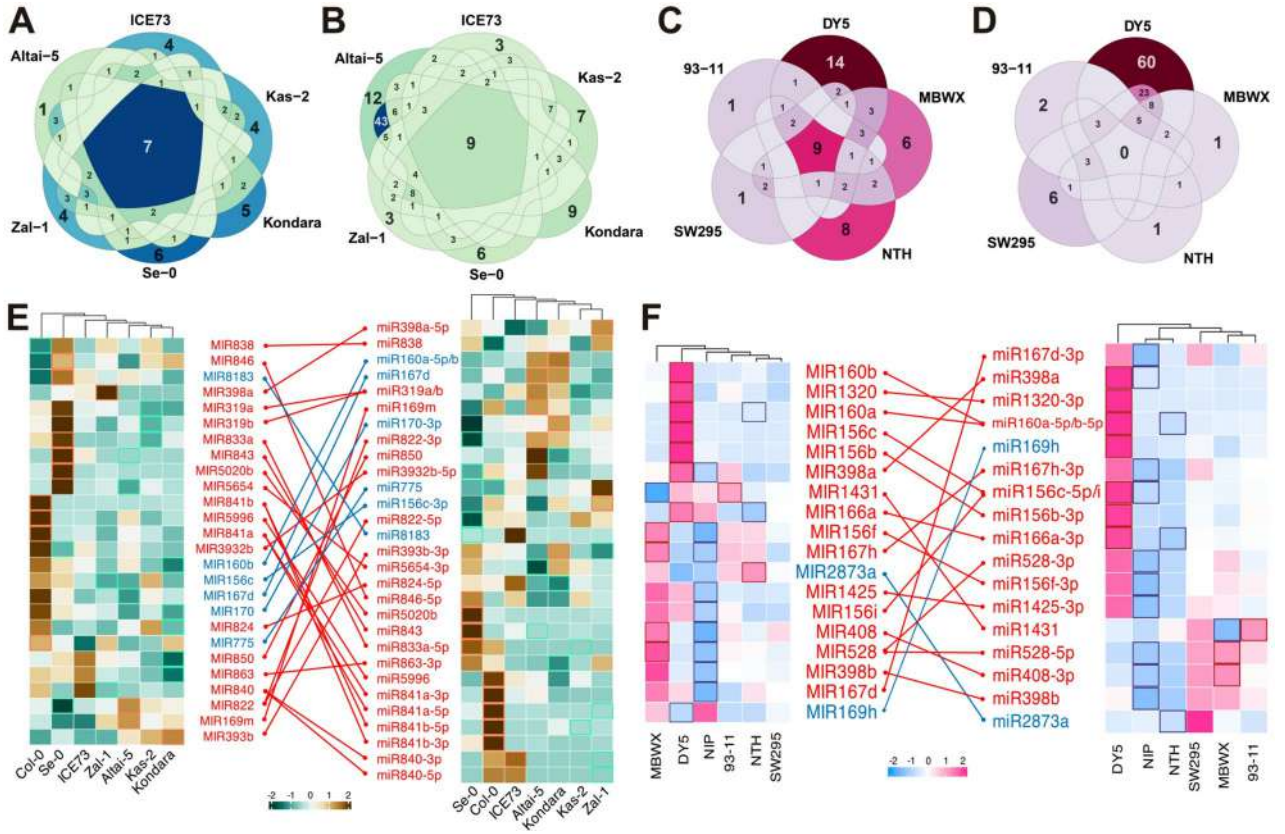


**Fig. 1** Differential expression analysis of the microRNA (miRNA) precursors and the mature miRNAs in distinct *Arabidopsis* accessions and rice varieties. **A** Venn diagram showing the differentially expressed miRNA precursors identified from the six accessions of *Arabidopsis*, as compared to Col-0 ($|\log_2 FC| \geq 1$, FDR < 0.05). **B** Venn diagram showing the differentially expressed mature miRNAs identified from the six *Arabidopsis* accessions ($|\log_2 FC| \geq 1$, FDR < 0.05). **C** Venn diagram showing the differentially expressed miRNA precursors identified from the five varieties of rice, as compared to Nipponbare ($|\log_2 FC| \geq 1$, FDR < 0.05). **D** Venn diagram showing the differentially expressed mature miRNAs identified from the five rice varieties ($|\log_2 FC| \geq 1$, FDR < 0.05). **E** Expression relationship between the miRNA precursors and the mature miRNAs in *Arabidopsis*. The differentially expressed precursors and mature miRNAs as shown in (A) and (B) were included in this analysis. We focused on the accession(s) in which a miRNA precursor was expressed at the highest or lowest level (highlighted with a red or green border line respectively in the heatmap on the left). We also focused on the accession(s) in which a mature miRNA was expressed at the highest or lowest level (highlighted with a red or green border line respectively in the heatmap on the right). If a consistent expression pattern was observed between the miRNA precursor and its mature miRNA, they would be highlighted in red and connected by a red line. If an opposite expression pattern was observed, they would be highlighted in blue and connected by a blue line. **F** Expression relationship between the miRNA precursors and the mature miRNAs in rice. A similar analysis was performed as described in (E). The heatmaps in (E)and (F) were drawn by using the ComplexHeatmap package [20]

miRNAs in Altai-5, ICE73, Kas-2, Kondara, Se-0 and Zal-1, respectively ($|\log_2 FC| \geq 1$, FDR < 0.05; Fig. 1B and Table S3). Notably, nine mature miRNAs are differentially expressed in all of the six accessions investigated. Among the nine miRNAs, five (ath-miR397b, ath-miR5651, ath-miR1886.2, ath-miR841a-5p and ath-miR841a-3p) are uniformly down-regulated, and two (ath-miR391-5p and ath-miR842) are uniformly up-regulated in the six accessions (Table S3). Besides, 12, 3, 7, 9, 6 and 3 differentially expressed mature miRNAs were specifically discovered in Altai-5, ICE73, Kas-2, Kondara, Se-0 and Zal-1, respectively (Fig. 1B and Table S3). In rice, 738 mature miRNAs were registered in miRBase (release 22). Compared to NIP, our analysis uncovered 15 (12 up-regulated and 3 down-regulated), 101 (94 up-regulated and 7 down-regulated), 21 (19 up-regulated and 2 down-regulated), 8 (3 up-regulated and 5 down-regulated), 49 (46 up-regulated and 3 down-regulated) differentially expressed mature miRNAs in 93 – 11, DY5, MBWX, NTH and SW295, respectively ($|\log_2 FC| \geq 1$, FDR < 0.05; Fig. 1D and Table S4). Different from *Arabidopsis*, none of the differentially expressed miRNAs was shared by the five varieties investigated. Instead, among the differentially expressed mature miRNAs, osa-miR3980a-3p and osa-miR3980b-3p were specifically discovered in 93 – 11, osa-miR156j-3p was specifically discovered in MBWX, osa-miR396c-3p was specifically discovered in NTH, osa-miR5150-5p, osa-miR2880, osa-miR1863c, osa-miR166d-5p, osa-miR2873a and osa-miR1432-5p were specifically discovered in SW295. Interestingly, a total of 60 differentially expressed mature miRNAs were specifically discovered in DY5 (Fig. 1D and Table S4).

Then, we questioned the expression correlation between precursors and mature miRNAs. To this end, the differentially expressed precursors and mature miRNAs identified above (Table S1 to S4) were analyzed separately. A pairwise comparison was performed between the precursors and their mature miRNAs, in order to identify the precursor—mature miRNA pairs with consistent or opposite expression patterns. Compared to the other subspecies, if the precursor and its mature miRNA were expressed at the highest or lowest levels in the same subspecies, the "precursor—mature" pair was regarded to be consistently expressed. If the precursor was expressed at the highest level while its mature miRNA was expressed at the lowest level in the same subspecies, or vice versa, the "precursor—mature" pair was regarded to be oppositely expressed. As a result, 23 pairs involving 20 precursors and 23 mature miRNAs with consistent expression patterns, and six pairs involving six precursors and six mature miRNAs with opposite expression patterns were discovered in *Arabidopsis* (Fig. 1E and Table S5). For example, the precursor ath-MIR838 shares a consistent expression pattern with its mature miRNA ath-miR838, both of which are expressed at the lowest levels in Col-0. The precursor ath-MIR846 also shares a consistent expression pattern with its mature miRNA ath-miR846-5p, both of which are expressed at the highest levels in Se-0. On the other hand, the precursor ath-MIR8183 is expressed at the highest level, while its mature miRNA ath-miR8183 is expressed at the lowest level in Se-0. In rice, 17 pairs involving 16 precursors and 17 mature miRNAs with consistent expression patterns, and two pairs involving two precursors and two mature miRNAs with opposite expression patterns were discovered (Fig. 1F and Table S6). For example, the precursor osa-MIR1320 shares a consistent expression pattern with its mature miRNA osa-miR1320-3p, both of which are expressed at the highest levels in DY5. The precursor osa-MIR398a also shares a consistent expression pattern with its mature miRNA osa-miR398a, both of which are expressed at the highest levels in DY5 and at the lowest levels in NIP. On the other hand, the precursor osa-MIR2873a is expressed at the highest level, while its mature miRNA osa-miR2873a is expressed at the lowest level in NTH. Summarily, the above cases show that in both *Arabidopsis* and rice, the mature miRNAs usually share the consistent expression patterns with their precursors.

#### Differential m$^6$A modification of the miRNA precursors

In both plants and animals, m$^6$A modification has been reported to play an important role in miRNA biogenesis [10–13]. Hence, MeRIP-seq with three biological replicates was performed for each subspecies to identify the differential m$^6$A peaks. The software exomePeak2 [27, 28] was applied for m$^6$A peak calling and differential peak identification with the "whole-genome" mode. All of the differential peaks were identified under the cutoff "$|\log_2 FC| > 0.5$, $P_{adj} < 0.05$". Compared to Col-0, 7,997 and 6,500 differential m$^6$A peaks were discovered in Altai-5 and ICE73 respectively. And, 5,576 (69.73%) and 4,527 (69.65%) peaks were hyper-methylated in the two accessions respectively. A total of 5,762, 9,893, 4,536 and 5,377 differential peaks were discovered in Kas-2, Kondara, Se-0 and Zal-1 respectively. However, only 1,256 (21.80%), 1,022 (10.33%), 738 (16.27%) and 1,191 (22.15%) were hyper-methylated in the four accessions respectively (Fig. 2A and Table S7). Thus, the six *Arabidopsis* accessions can be classified into two groups: one group includes Altai-5 and ICE73, and the other group includes Kas-2, Kondara, Se-0 and Zal-1. Compared to NIP, a total of 7,990, 7,320, 7,111 and 7,257 differential m$^6$A peaks were discovered in 93 – 11, MBWX, NTH and SW295 respectively. And, 6,334 (79.27%), 4,925 (67.28%), 4,947 (69.57%) and 5,105 (70.35%) peaks were hyper-methylated in the four rice varieties respectively. In DY5, 7,933 differential peaks were discovered, and only 2,000 (25.21%) peaks were hyper-methylated (Fig. 2B and Table S8). Thus, the five rice varieties can be classified into two groups: one group includes 93 – 11, MBWX, NTH and SW295, and the other group includes DY5.

Next, we investigated the genomic distribution of the above identified peaks. The genomic regions were classified into "5'UTRs (untranslated regions)", "exons", "introns", "3'UTRs", "1-kb upstream regions of the protein-coding genes" and "1-kb downstream regions of the protein-coding genes". In *Arabidopsis*, quite distinct genomic distribution patterns were observed for the two groups defined above. Specifically, in Altai-5 and ICE73, more than 50% of the hyper-methylated peaks were assigned to "3'UTRs", and around 14% were assigned to "exons". However, in the other four accessions, less than 30% of the hyper-methylated peaks were assigned to "3'UTRs", and higher than 30% were assigned to "exons". Besides, in both Altai-5 and ICE73, around 50% of the hypo-methylated peaks were assigned to "3'UTRs", and more than 20% were assigned to "exons". However, in the other four accessions, more than 54% of the hypo-methylated peaks were assigned to "3'UTRs", and less than 17% were assigned to "exons" (Fig. 2C and Table S7). In rice, distinct genomic distribution patterns were also observed for the two groups defined above. Specifically, in DY5, 33% and 33.5% of the hyper-methylated peaks were assigned to "3'UTRs" and "exons" respectively, and about 60% and 15% of the hypo-methylated peaks were assigned to "3'UTRs" and "exons" respectively. However, in the other four varieties, more than 54% of the hyper-methylated peaks were assigned to "3'UTRs", and less than 15% were assigned to "exons". And, less than 43% of

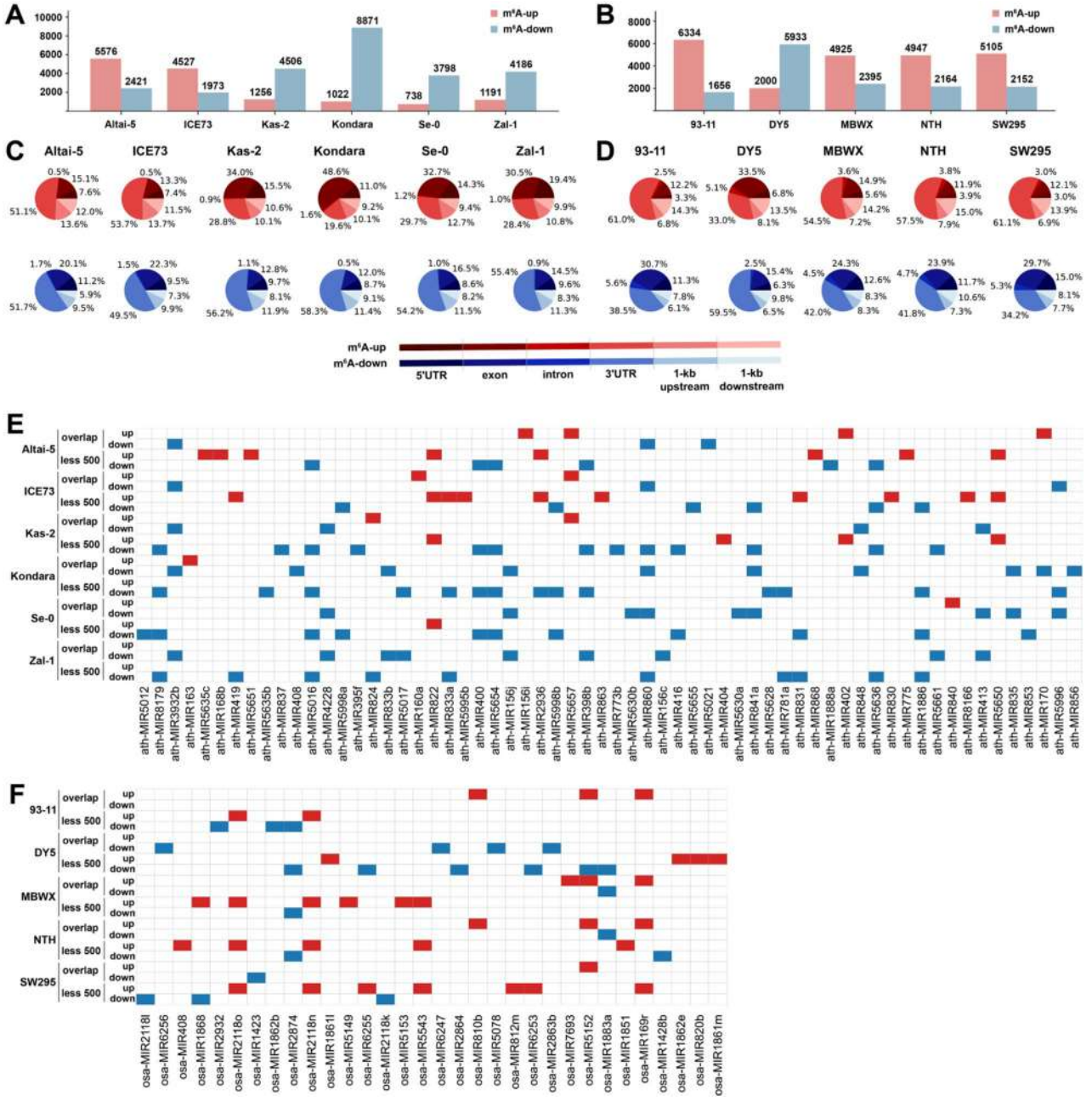


**Fig. 2** Transcriptome-wide differential m$^6$A methylation among distinct *Arabidopsis* accessions and rice varieties, and the microRNAs (miRNAs) associated with the differential m$^6$A peaks in each accession or variety. **A** Numbers of the differential m$^6$A peaks identified from the six accessions of *Arabidopsis*. For each accession, Col-0 was treated as a control for the differential peak identification. **B** Numbers of the differential m$^6$A peaks identified from the five varieties of rice. For each accession, NIP was treated as a control for the differential peak identification. **C** Genome-wide distribution patterns of the differential m$^6$A peaks identified from the six accessions of *Arabidopsis*. The peaks were assigned to 5′UTRs, exons, introns, 3′UTRs, or the 1-kb regions upstream or downstream of the protein-coding genes. **D** Genome-wide distribution patterns of the differential m$^6$A peaks identified from the five varieties of rice. The peaks were assigned to 5′UTRs, exons, introns, 3′UTRs, or the 1-kb regions upstream or downstream of the protein-coding genes. From (A) to (D), the differential peaks were identified under the cutoff "$|\log_2 FC| > 0.5$, $P_{adj} < 0.05$", and the hyper- and hypo-methylated peaks were represented by red and blue respectively. **E** The miRNA precursors associated with the differential m$^6$A peaks in each *Arabidopsis* accession. **F** The miRNA precursors associated with the differential m$^6$A peaks in each rice variety. For both (E) and (F), the differential peaks were classified into two major types: the peaks overlapped with the miRNA precursors, and the peaks adjacent to the miRNA precursors with a distance less than 500 nt. Each major type was further divided into hyper-methylated peaks (cells filled in red) and hypo-methylated peaks (cells filled in blue)

the hypo-methylated peaks were assigned to "3'UTRs", and more than 23% were assigned to "exons" (Fig. 2D and Table S8). Taken together, a large portion of the differential m$^6$A peaks were assigned to "3'UTRs", while much less were assigned to "5'UTRs", which was consistent with the reported patterns in plants [33]. However, subtle but distinguishable differences were observed among different subspecies.

To investigate the potential role of m$^6$A modification in differential miRNA expression, a genome-wide search was performed to identify the miRNA precursors associated with the differential m$^6$A peaks in *Arabidopsis* and rice. The genomic information of the miRNA precursors was obtained from miRBase (release 22) [32]. Since the miRBase-registered precursors are shorter than the primary miRNAs (pri-miRNAs) in many cases, the miRNA-associated peaks were divided into two classes: (1) the peaks overlapped with the miRBase-registered precursors; and (2) the peaks overlapped with the 500-nt upstream or downstream regions of the precursors. As a result, in *Arabidopsis*, 21, 21, 23, 25, 22 and 19 miRNA precursors were discovered to be associated with the differential m$^6$A peaks in Altai-5, ICE73, Kas-2, Kondara, Se-0 and Zal-1, respectively (Fig. 2E and Table S9). Interestingly, for the last four accessions, most of the miRNA-associated m$^6$A peaks are hypo-methylated. In rice, a total of 8, 14, 11, 11 and 12 miRNA precursors were discovered to be associated with the differential m$^6$A peaks in 93 – 11, DY5, MBWX, NTH and SW295, respectively (Fig. 2F and Table S10). Notably, except for DY5, most of the miRNA-associated m$^6$A peaks are hyper-methylated.

Since the methylated "A" sites could be interfered by genomic sequence polymorphisms which will lead to differential m$^6$A modification among distinct subspecies, the genomic polymorphism data of *Arabidopsis* and rice were retrieved from the *Arabidopsis* 1001 Genomes Project [1] and the Rice 3,000 Genomes Project [2] respectively. Then, we searched for the polymorphic sites resided within the differential m$^6$A peaks associated with the miRNA precursors. Compared to the reference subspecies (Col-0 and NIP), the polymorphic sites were classified into two types: (I) other nucleotides to "A", which may create an m$^6$A site resulting in hyper-methylation; and (II) "A" to other nucleotides, which may destroy an m$^6$A site resulting in hypo-methylation. As a result, 20, 12, 14, 2 and 8 type I polymorphic sites were discovered within the hyper-methylated peaks associated with six, three, two, one and one miRNA precursors in Altai-5, ICE73, Kas-2, Kondara and Se-0, respectively (highlighted in red in Table S9). And, 10, 4, 24, 17, 23 and 13 type II polymorphic sites were discovered within the hypo-methylated peaks associated with five, two, six, six, nine and six miRNA precursors in Altai-5, ICE73, Kas-2, Kondara, Se-0 and Zal-1, respectively (highlighted in blue in Table S9). However, in rice, only one type II polymorphic site was discovered within the hypo-methylated peak associated with osa-MIR2863b in DY5, and another type II polymorphic site was discovered within the hypo-methylated peak associated with osa-MIR1428b in NTH (highlighted in blue in Table S10).

### Differential m$^6$A modification related to differential miRNA expression

The above analysis showed that dozens of the miRNA precursors were associated with the differential m$^6$A peaks. On the other hand, m$^6$A modification has been reported to have a significant influence on transcript stability [34]. Thus, it will be interesting to investigate whether differential m$^6$A modification can lead to the differential expression of the miRNA precursors as observed above (Fig. 1). Indeed, in *Arabidopsis*, five (ath-MIR170, ath-MIR2936, ath-MIR3932b, ath-MIR5021 and ath-MIR5654), four (ath-MIR833a, ath-MIR841a, ath-MIR3932b and ath-MIR5996), four (ath-MIR413, ath-MIR841a, ath-MIR3932b and ath-MIR5661), seven (ath-MIR156j, ath-MIR170, ath-MIR841a, ath-MIR2936, ath-MIR3932b, ath-MIR5628 and ath-MIR5996), six (ath-MIR413, ath-MIR822, ath-MIR840, ath-MIR841a, ath-MIR5654 and ath-MIR5996) and four (ath-MIR156c, ath-MIR413, ath-MIR833a and ath-MIR3932b) precursors associated with the differential m$^6$A peaks are differentially expressed in Altai-5, ICE73, Kas-2, Kondara, Se-0 and Zal-1, respectively ($|\log_2 FC| \geq 1$, FDR < 0.05; Table S9). In rice, osa-MIR2932 in 93 – 11, osa-MIR820b, osa-MIR2863b and osa-MIR6247 in DY5, osa-MIR1883a and osa-MIR7693 in MBWX, and osa-MIR408 in NTH are associated with the differential m$^6$A peaks and differentially expressed ($|\log_2 FC| \geq 1$, FDR < 0.05; Table S10). It has also been reported that m$^6$A modification plays an important regulatory role in miRNA processing in both plants and animals [10–13]. In this regard, a comprehensive search was performed to identify the differentially expressed mature miRNAs whose precursors were associated with the differential m$^6$A peaks. As a result, six (ath-miR168b-3p, ath-miR168b-5p, ath-miR398b-5p, ath-miR400, ath-miR5651 and ath-miR5654-3p), four (ath-miR833a-5p, ath-miR841a-3p, ath-miR841a-5p and ath-miR1886.2), four (ath-miR398b-3p, ath-miR841a-3p, ath-miR841a-5p and ath-miR3932b-5p), 13 (ath-miR163, ath-miR398b-5p, ath-miR400, ath-miR408-3p, ath-miR781a, ath-miR833a-5p, ath-miR841a-3p, ath-miR841a-5p, ath-miR848, ath-miR1886.1, ath-miR1886.2, ath-miR3932b-5p and ath-miR5996), nine (ath-miR400, ath-miR822-3p, ath-miR822-5p, ath-miR840-3p, ath-miR840-5p, ath-miR841a-3p, ath-miR841a-5p, ath-miR1886.2 and ath-miR5996), and six (ath-miR1886.1, ath-miR1886.2, ath-miR3932b-5p, ath-miR398b-5p, ath-miR833a-5p and ath-miR833b)

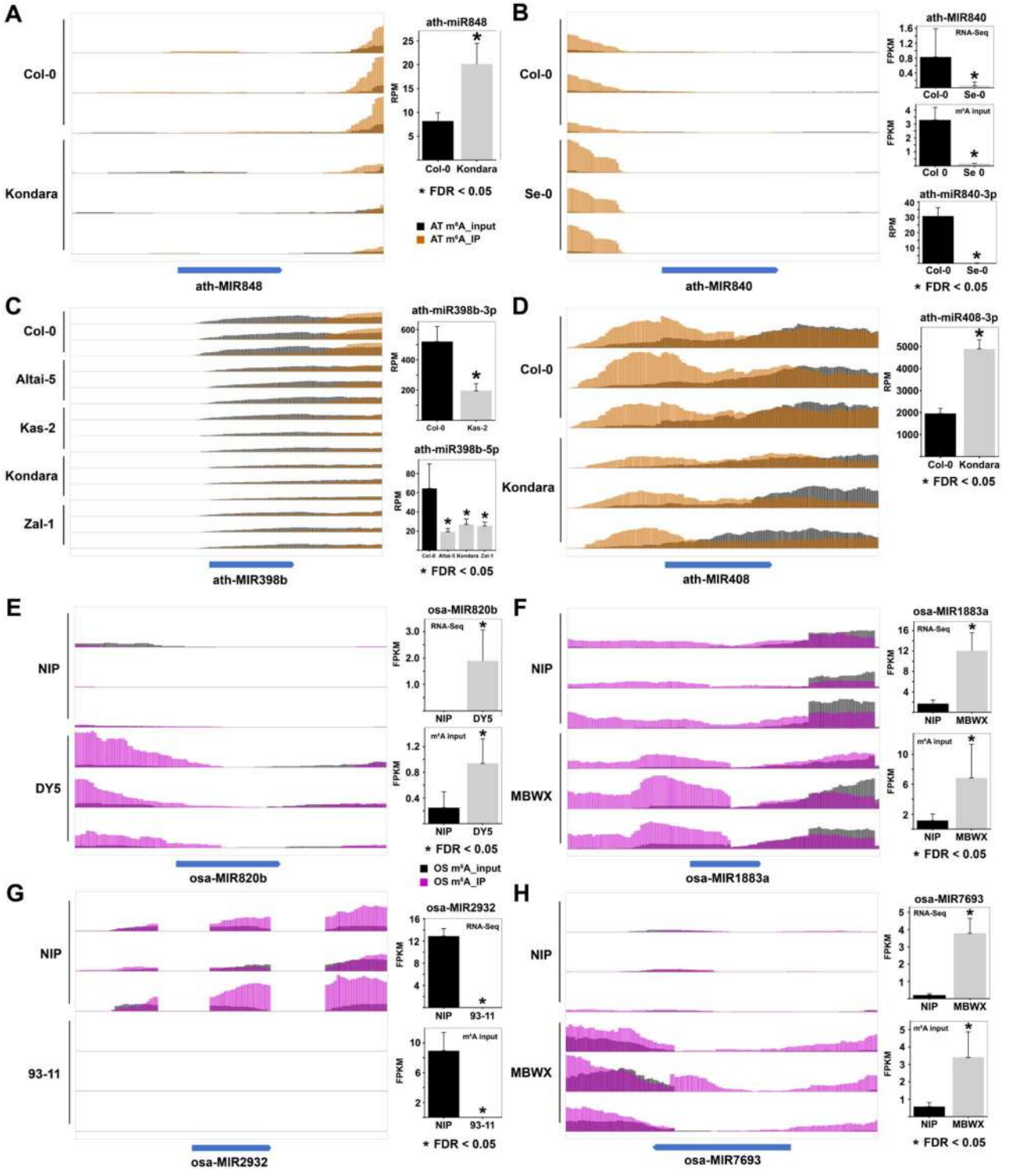


**Fig. 3** (See legend on next page.)

mature miRNAs are differentially expressed in Altai-5, ICE73, Kas-2, Kondara, Se-0 and Zal-1, respectively ($|\log_2 FC| \geq 1$, FDR < 0.05; Table S9). In rice, osa-miR1861m in DY5 and osa-miR812m in SW295 were identified to be differentially expressed ($|\log_2 FC| \geq 1$, FDR < 0.05; Table S10).

The above results indicate that the differential $m^6A$ modification may have a direct impact on precursor stability or miRNA processing. By using the IGV

(See figure on previous page.)

**Fig. 3** Differential microRNA (miRNA) expression potentially resulted from the differential m$^6$A modification among distinct *Arabidopsis* accessions or rice varieties. **A** Compared to Col-0, a hypo-methylated peak adjacent to the precursor ath-MIR848 was detected in Kondara. As shown in the histogram, the level of the mature miR848 is significantly higher in Kondara. **B** Compared to Col-0, a hyper-methylated peak adjacent to the precursor ath-MIR840 was detected in Se-0. As shown in the histograms, the levels of both the precursor MIR840 and the mature miR840-3p are dramatically reduced in Se-0. **C** Compared to Col-0, the hypo-methylated peaks overlapped with the precursor ath-MIR398b were detected in Altai-5, Kas-2, Kondara and Zal-1. As shown in the histograms, the level of the mature miR398b-3p is significantly lower in Kas-2, and the levels of the mature miR398b-5p are significantly lower in Altai-5, Kondara and Zal-1. **D** Compared to Col-0, a hypo-methylated peak overlapped with the precursor ath-MIR408 was detected in Kondara. As shown in the histogram, the level of the mature miR408-3p is significantly higher in Kondara. **E** Compared to NIP, a hyper-methylated peak overlapped with the precursor osa-MIR820b was detected in DY5. As shown in the histogram, the level of the precursor MIR820b is significantly up-regulated in DY5. **F** Compared to NIP, a hyper-methylated peak overlapped with the precursor osa-MIR1883a was detected in MBWX. As shown in the histogram, the level of the precursor MIR1883a is significantly up-regulated in MBWX. **G** Compared to NIP, a hypo-methylated peak overlapped with the precursor osa-MIR2932 was detected in 93 – 11. As shown in the histogram, the level of the precursor MIR2932 is dramatically down-regulated in 93 – 11. **H** Compared to NIP, a hyper-methylated peak overlapped with the precursor osa-MIR7693 was detected in MBWX. As shown in the histogram, the level of the precursor MIR7693 is significantly up-regulated in MBWX. For all panels, the IGV software [29] was used for m$^6$A peak visualization, and the miRNA precursors are represented by the blue horizontal bars. The MeRIP-seq data of three biological replicates are presented. The expression patterns of the miRNA precursors are supported by both the RNA-seq data and the "input" data of MeRIP-seq

visualization tool [29], some typical examples are presented here for detailed illustration. Compared to Col-0, a hypo-methylated m$^6$A peak close to ath-MIR848 was detected in Kondara. On the other hand, the level of the mature miRNA ath-miR848 in Kondara is significantly higher than that in Col-0 (FDR < 0.05; Fig. 3A). In Se-0, a hyper-methylated m$^6$A peak close to the precursor ath-MIR840 was discovered. Notably, in addition to the mature miRNA ath-miR840-3p, the level of ath-MIR840 is significantly lower in Se-0 than in Col-0, which is supported by both the RNA-seq data and the "input" data of MeRIP-seq (FDR < 0.05; Fig. 3B). For another case of *Arabidopsis*, the m$^6$A peak overlapped with ath-MIR398b is hypo-methylated in Altai-5, Kas-2, Kondara and Zal-1. Compared to Col-0, the level of the mature miRNA ath-miR398b-3p is significantly lower in Kas-2, and the levels of the other mature miRNA ath-miR398b-5p are significantly lower in Altai-5, Kondara and Zal-1 (FDR < 0.05; Fig. 3C). Another m$^6$A peak overlapped with ath-MIR408 is hypo-methylated in Kondara, and the level of the mature miRNA ath-miR408-3p is significantly higher in this accession compared to Col-0 (FDR < 0.05; Fig. 3D). In rice, the m$^6$A peak overlapped with the precursor osa-MIR820b is hyper-methylated in DY5, and the level of osa-MIR820b is significantly higher in this variety compared to NIP (FDR < 0.05; Fig. 3E). In the rice variety MBWX, the m$^6$A peak overlapped with osa-MIR1883a is hyper-methylated. Accordingly, the level of this precursor is significantly higher in MBWX (FDR < 0.05; Fig. 3F). For another case of rice, three m$^6$A peaks overlapped with or close to osa-MIR2932 were discovered in NIP. However, the m$^6$A modification signal could be hardly detected in the variety 93 – 11. Accordingly, the level of this precursor could also be hardly detected in 93 – 11 (FDR < 0.05; Fig. 3G). Another m$^6$A peak overlapped with osa-MIR7693 is hyper-methylated in MBWX. Accordingly, the level of osa-MIR7693 is significantly higher in this variety (FDR < 0.05; Fig. 3H). Notably, in the four cases of rice, the expression patterns of the miRNA precursors are supported by both the RNA-seq data and the "input" data of MeRIP-seq. Taken together, our results indicate that differential m$^6$A modification has a notable influence on miRNA expression likely through the regulation of precursor stability and/or miRNA maturation.

### Genomic polymorphisms related to differential miRNA expression

Since the hairpin-like structures of the miRNA precursors are indispensible for their efficient processing [3], the DNA polymorphism-induced structure variations have a reasonable impact on miRNA maturation. In this regard, the genomic polymorphism data of *Arabidopsis* and rice, mainly including SNPs, were retrieved from the *Arabidopsis* 1001 Genomes Project [1] and the Rice 3,000 Genomes Project [2] respectively. Then, we searched for the polymorphic sites on the miRNA precursors, and evaluated their effects on the precursor structures based on the RNAfold predictions [24]. In *Arabidopsis*, referring to Col-0, two SNP sites were identified on the mature miRNA-coding region of the precursor ath-MIR833a in Altai-5, ICE73, Kondara and Zal-1 respectively, and one SNP site was allocated to the mature miRNA-coding region in Kas-2 (Fig. 4A). In each case, the sequence complementarity of the stem region encoding the mature miRNA ath-miR833a-5p is obviously reduced by the SNPs. Notably, the precursor levels are decreased in the five accessions, especially in ICE73, Kas-2 and Zal-1 (FDR < 0.05). This result is supported by both the RNA-seq data and the "input" data of MeRIP-seq (Fig. 4A). Besides, the levels of ath-miR833a-5p could be hardly detected in all of the five accessions (Fig. 4A), indicating that the miRNA processing was significantly inhibited. Also in *Arabidopsis*, three SNP sites were discovered on the precursor ath-MIR393b in Altai-5 and Kas-2 respectively, and two SNP sites were discovered in Zal-1. In each accession, a large loop at the bottom of the mature miRNA-coding region is introduced by a SNP site, which may interfere efficient miRNA

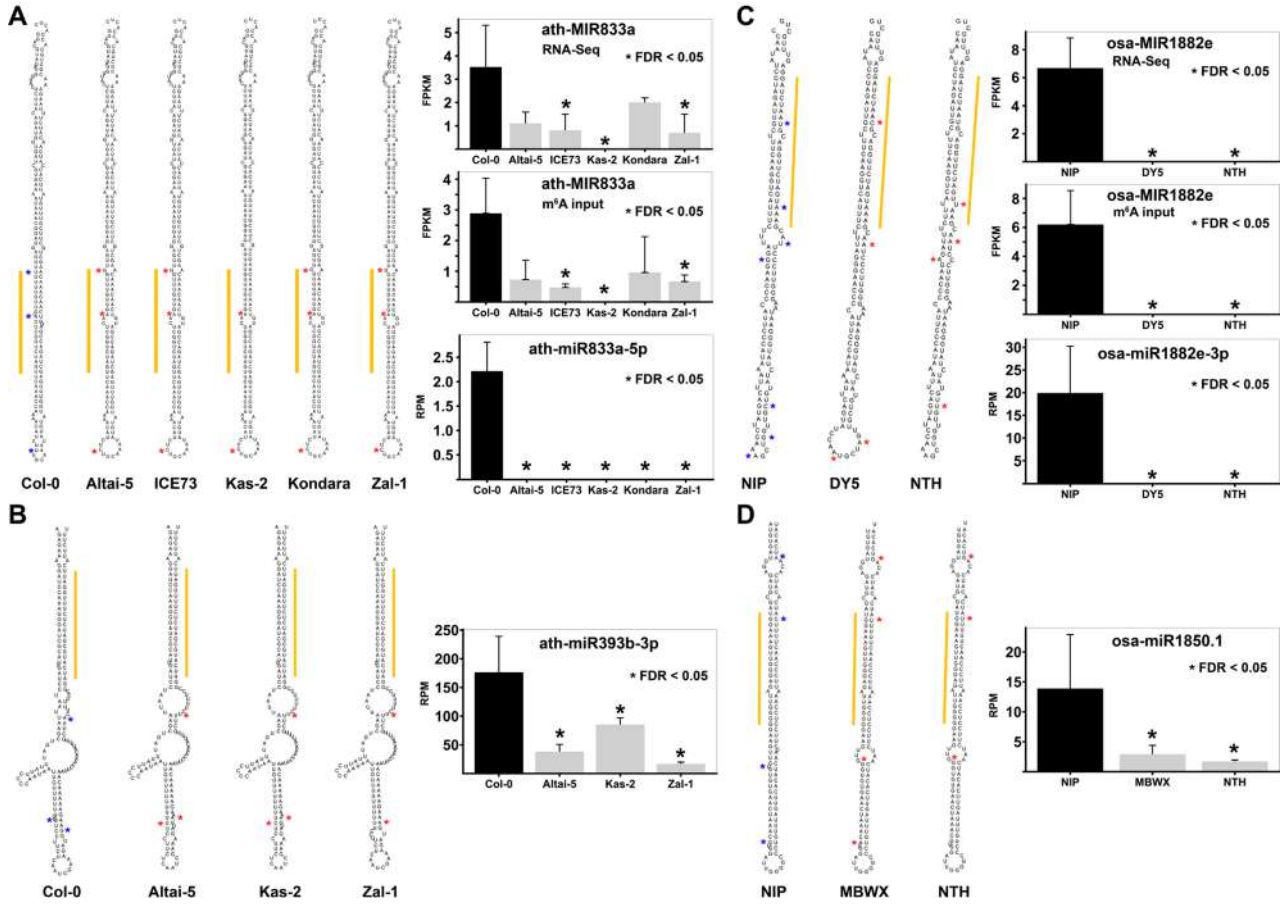


**Fig. 4** Effects of the single nucleotide polymorphisms (SNPs) on microRNA (miRNA) structures and expression. **A** The secondary structure and expression of ath-MIR833a affected by the SNPs. The differential expression levels of the precursors ath-MIR833a and the mature miRNA ath-miR833a-3p are shown by the histograms on the right. **B** The secondary structure and expression of ath-MIR393b affected by the SNPs. The differential expression level of the mature miRNA ath-miR393b-3p is shown by the histogram on the right. For (A) and (B), the SNP sites on the miRNA precursors are marked by the blue and red asterisks in the reference accession Col-0 and the other accessions respectively. The mature miRNA-coding regions are marked by yellow lines on the precursors. **C** The secondary structure and expression of osa-MIR1882e affected by the SNPs. The differential expression levels of the precursors osa-MIR1882e and the mature miRNA osa-miR1882e-3p are shown by the histograms on the right. **D** The secondary structure and expression of osa-MIR1850 affected by the SNPs. The differential expression level of the mature miRNA osa-miR1850.1 is shown by the histogram on the right. For (C) and (D), the SNP sites on the precursors are marked by the blue and red asterisks in the reference variety NIP and the other varieties respectively. The mature miRNA-coding regions are marked by the yellow lines on the precursors. The secondary structures of the miRNA precursors were predicted by using RNAfold [24]. The expression patterns of the miRNA precursors are supported by both the RNA-seq data and the "input" data of MeRIP-seq

processing. Indeed, the mature miRNA ath-miR393b-3p is significantly repressed in all of the three accessions, as compared to Col-0 (FDR < 0.05; Fig. 4B). In rice, referring to NIP, four SNP sites were identified on the precursor osa-MIR1882e in DY5 and NTH respectively. The secondary structures of osa-MIR1882e are significantly altered. Accordingly, the precursor levels are greatly reduced in the two varieties, which is supported by the RNA-seq data and the "input" data of MeRIP-seq. The accumulation of the mature miRNA osa-miR1882e-3p is also greatly repressed in both varieties (FDR < 0.05; Fig. 4C). Also in rice, several SNP sites were identified on the precursor osa-MIR1850 in MBWX and NTH. In each variety, a loop structure at the bottom of the mature miRNA-coding region is introduced by a SNP site, which may interfere efficient miRNA processing. Indeed, the levels of the mature miRNA osa-miR1850.1 are significantly reduced in both varieties (FDR < 0.05; Fig. 4D). Some additional cases were discovered for ath-MIR172e, ath-MIR777, ath-MIR822 and ath-MIR833b in *Arabidopsis*, and osa-MIR156g in rice (Fig. S2). Summarily, the SNP sites identified above have a direct impact on the structures of the miRNA precursors, thus likely interfering the stability and processing of the precursors.

## Target gene regulation affected by differential miRNA expression

The above results indicate that both genomic and epitranscriptomic variations have potential impacts on differential miRNA expression. However, the biological relevance

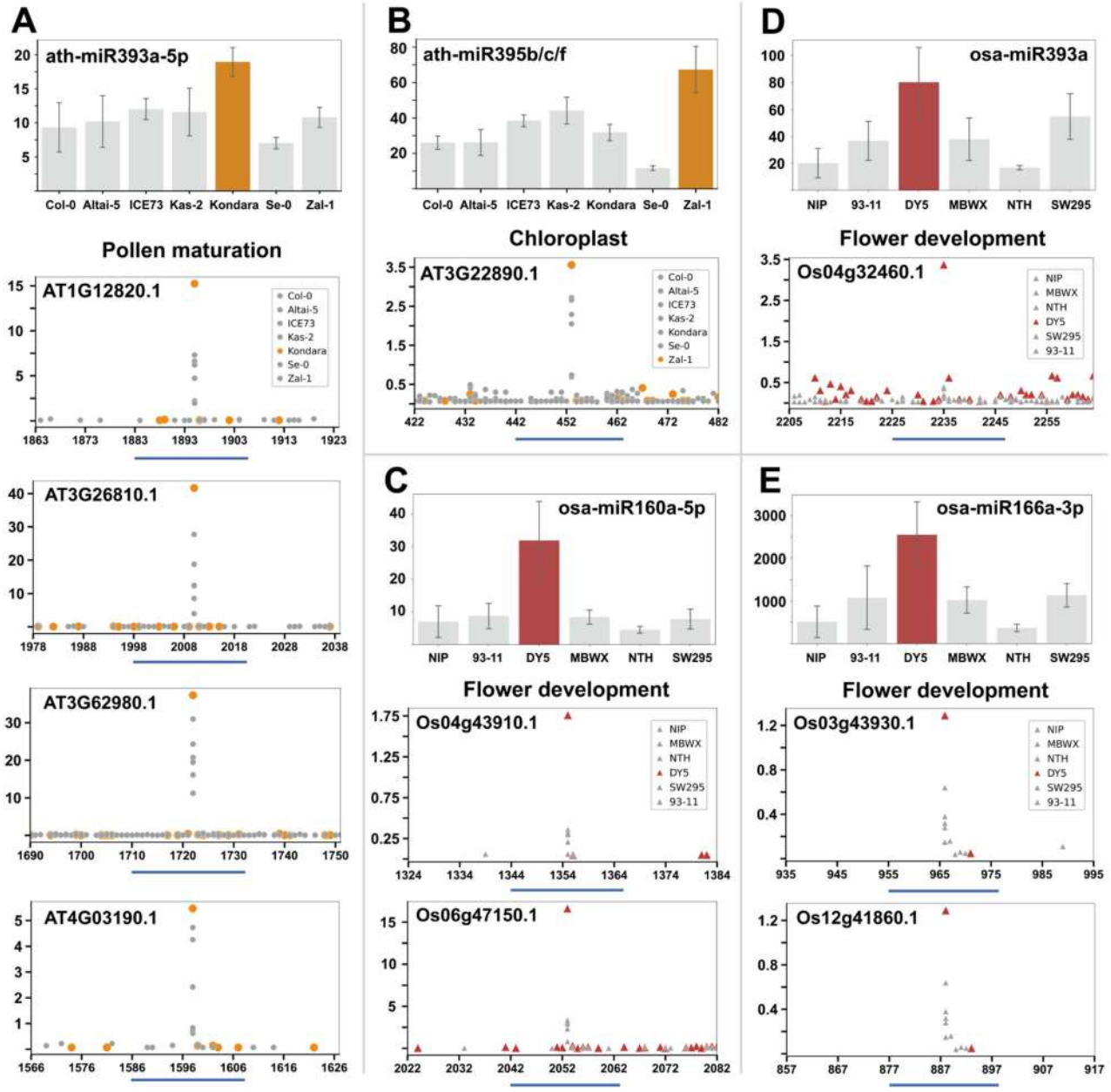


**Fig. 5** (See legend on next page.)

remains unclear. In plants, miRNAs regulate gene expression mostly through site-specific transcript cleavages [3]. Degradome-seq is a powerful tool for miRNA target validation, and the regulatory intensity can be evaluated based on the signature levels [5, 6]. In this regard, the degradome-seq experiment was performed for the seven accessions of *Arabidopsis* and the six varieties of rice, in order to measure the regulatory intensity in RPM. On the other hand, among the differentially expressed miRNAs identified above (Fig. 1B and D), we focused on the highly expressed miRNAs specifically present in one subspecies. The miRU algorithm [22, 23] was employed to predict the miRNA targets. Then, the predicted miRNA—target pairs were subject to degradome-based validation as described in the previous studies [5, 6]. For each validated pair, the intensity of the site-specific cleavage signals was compared among different subspecies based on the levels of the degradome signatures. Only those pairs with cleavage signals well correlated with the subspecies-specific expression patterns of the regulatory miRNAs were retained for functional analysis.

As a result, several miRNA—target pairs with notable biological functions were identified. In *Arabidopsis*, the mature miRNA ath-miR393a-5p is highly expressed in

(See figure on previous page.)

**Fig. 5** The correlation between the mature microRNA (miRNA) abundance and its regulatory intensity as exemplified by the degradome-based comparative analysis. **A** The high level of ath-miR393a-5p is well correlated to the enhanced cleavage signals on its target transcripts in Kondara. The histogram shows the levels of ath-miR393a-5p in different *Arabidopsis* accessions (measured by the *y* axis in RPM). The target plots (t-plots) show the degradome signals surrounding the miRNA binding sites (MBSs, indicated by the blue lines) on the targets. The *y* axes measure the degradome signal intensity in RPM, and the signals detected from Kondara were denoted by the yellow dots. All of the target genes are functionally related to pollen maturation. **B** The high level of ath-miR395b/c/f is well correlated to the enhanced cleavage signal on its target transcript in Zal-1. The histogram shows the levels of ath-miR395b/c/f in different *Arabidopsis* accessions (measured by the *y* axis in RPM). The t-plot shows the degradome signals surrounding the MBS (indicated by the blue line) on the target. The *y* axis measures the degradome signal intensity in RPM, and the signals detected from Zal-1 were denoted by the yellow dots. The target gene is functionally related to chloroplast. **C** The high level of osa-miR160a-5p is well correlated to the enhanced cleavage signals on its target transcripts in DY5. The histogram shows the levels of osa-miR160a-5p in different rice varieties (measured by the *y* axis in RPM). The t-plots show the degradome signals surrounding the MBSs (indicated by the blue lines) on the targets. The *y* axes measure the degradome signal intensity in RPM, and the signals detected from DY5 were denoted by the red triangles. The target genes are functionally related to flower development. **D** The high level of osa-miR393a is well correlated to the enhanced cleavage signal on its target transcript in DY5. The histogram shows the levels of osa-miR393a in different rice varieties (measured by the *y* axis in RPM). The t-plot shows the degradome signals surrounding the MBS (indicated by the blue line) on the target. The *y* axis measures the degradome signal intensity in RPM, and the signals detected from DY5 were denoted by the red triangles. The target gene is functionally related to flower development. **E** The high level of osa-miR166a-3p is well correlated to the enhanced cleavage signals on its target transcripts in DY5. The histogram shows the levels of osa-miR166a-3p in different rice varieties (measured by the *y* axis in RPM). The t-plots show the degradome signals surrounding the MBSs (indicated by the blue lines) on the targets. The *y* axes measure the degradome signal intensity in RPM, and the signals detected from DY5 were denoted by the red triangles. The target genes are functionally related to flower development. For all t-plots, the strong degradome signatures mapped to the middle of the MBSs were identified to be the cleavage signals

Kondara compared to the other six accessions. Accordingly, the cleavage signals detected on the four targets, AT1G12820.1, AT3G26810.1, AT3G62980.1 and AT4G03190.1, are specifically stronger in Kondara (Fig. 5A). Intriguingly, all of the target genes are functionally related to pollen maturation. To seek for the experimental evidence supporting the potential function, the gene expression data publicly available at the electronic Fluorescent Pictograph (eFP) Browser (https://www.bar.utoronto.ca/) [35] were used. By comparing the data covering 47 different tissues, the four genes, especially for *AT1G12820*, *AT3G62980* and *AT4G03190*, were observed to be expressed at quite low levels in mature pollen (Fig. S3A), pointing to the potentially negative role of these genes in pollen maturation. For another case of *Arabidopsis*, ath-miR395b, ath-miR395c and ath-miR395f, all of which shared the identical sequence, are highly expressed in Zal-1 compared to the other six accessions. Accordingly, the cleavage signals detected on the target AT3G22890.1 are specifically stronger in Zal-1 (Fig. 5B). The target gene is functionally related to chloroplast. Indeed, compared to the expression data of pollen, flowers, seeds and roots, *AT3G22890* is expressed at higher levels in the organs enriched with chloroplasts, including cotyledons, cauline leaves and rosette leaves. Interestingly, the gene expression could be induced by osmotic, drought and heat treatments respectively (Fig. S3B). In rice, the level of osa-miR160a-5p in DY5 is especially higher than those in the other five varieties. Accordingly, the cleavage signals detected on the two targets, LOC_Os04g43910.1 and LOC_Os06g47150.1, are specifically stronger in DY5 (Fig. 5C). The two target genes were annotated to be functionally involved in flower development. Indeed, the expression data retrieved from RGAP7.0 (https://rice.uga.edu/) [26] show that both genes are expressed at higher levels in inflorescences, as compared to the levels in seeds, seedlings, stems, leaves and roots. Besides, the expression levels of the two genes were elevated under the heat and cold stress treatments respectively (Fig. S3C). For another case of rice, osa-miR393a is highly expressed in DY5, and the cleavage signals detected on its target LOC_Os04g32460.1 are also specifically stronger in DY5 compared to the other five varieties (Fig. 5D). Supporting the annotated role of *LOC_Os04g32460* in flower development, the expression level of this gene is especially high in inflorescences, as compared to the other tissues. Besides, the gene expression was obviously repressed under the drought treatment (Fig. S3D). The mature miRNA osa-miR166a-3p is highly expressed in DY5, and the cleavage signals detected on its targets LOC_Os03g43930.1 and LOC_Os12g41860.1 are also specifically stronger in DY5 compared to the other varieties (Fig. 5E). Supporting the annotated role of *LOC_Os12g41860* in flower development, the expression level of this gene is especially high in inflorescences. Besides, the expression of the two target genes could be significantly induced by the cold treatment (Fig. S3E). Taken together, in both *Arabidopsis* and rice, the expression of most target genes could be affected by stress treatments, indicating their role in abiotic stress responses.

### MiRNA—target interactions interfered by DNA polymorphisms

The above results show that the subspecies-specific intensive target regulation may result from the well-correlated expression patterns of the regulatory miRNAs. On the other hand, the miRNA—target interactions can be interfered by sequence polymorphisms through a direct impact on the miRNA binding affinity. In this regard, by using the miRU algorithm [22, 23], all transcripts available in TAIR10 and RGAP7.0 were subject to

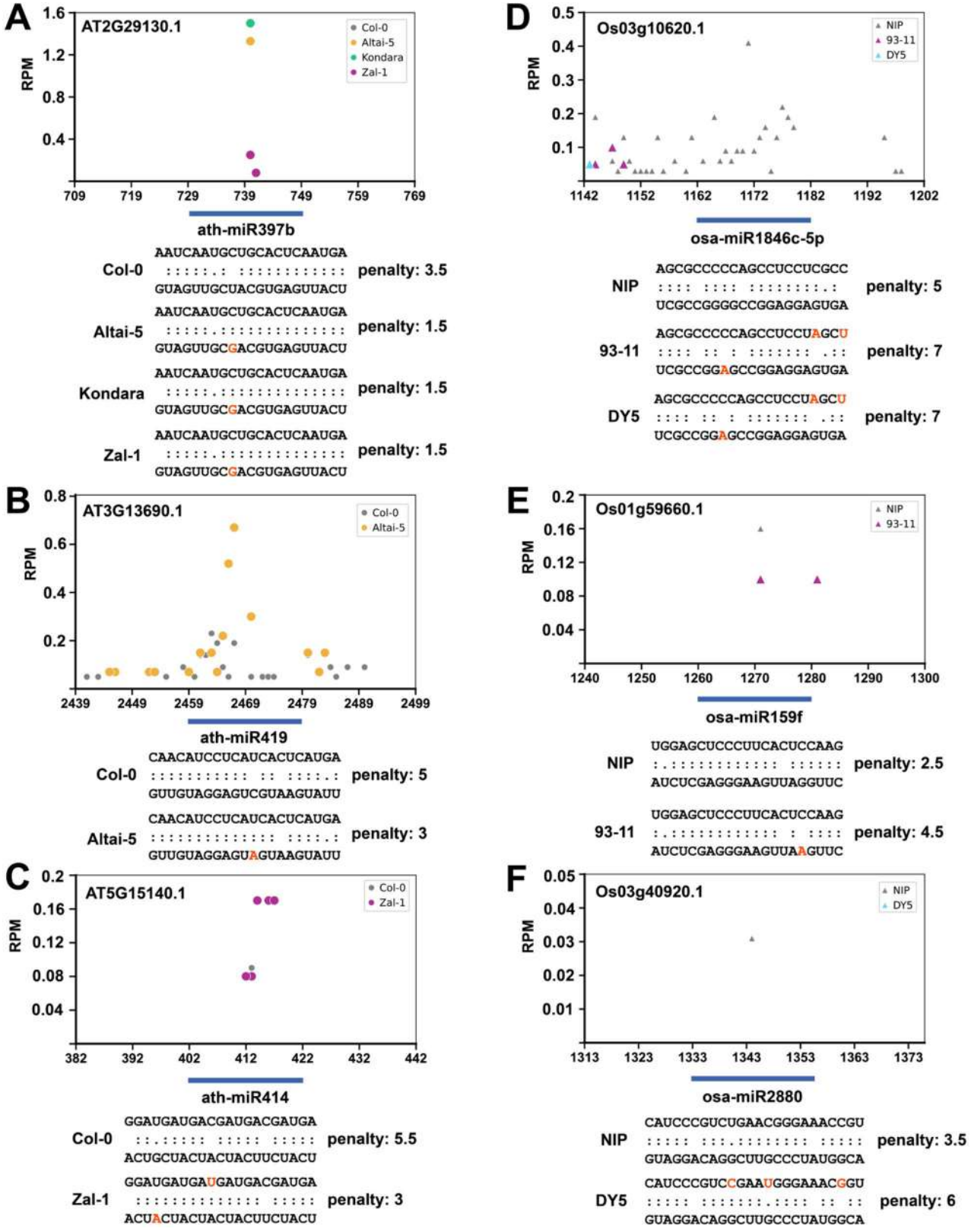


**Fig. 6** (See legend on next page.)

(See figure on previous page.)

**Fig. 6** Effects of the sequence polymorphisms on microRNA (miRNA)-mediated target regulation. **A** The impact of single nucleotide polymorphisms (SNPs) on differential regulation of the transcript AT2G29130.1 by ath-miR397b among the *Arabidopsis* accessions Col-0, Altai-5, Kondara and Zal-1. **B** The impact of SNPs on differential regulation of the transcript AT3G13690.1 by ath-miR419 between Col-0 and Altai-5. **C** The impact of SNPs on differential regulation of the transcript AT5G15140.1 by ath-miR414 between Col-0 and Zal-1. **D** The impact of SNPs on differential regulation of the transcript LOC_Os03g10620.1 by osa-miR1846c-5p among NIP, 93 – 11 and DY5. **E** The impact of SNPs on differential regulation of the transcript LOC_Os01g59660.1 by osa-miR159f between the rice varieties NIP and 93 – 11. **F** The impact of SNPs on differential regulation of the transcript LOC_Os03g40920.1 by osa-miR2880 between the rice varieties NIP and DY5. For all figure panels, the target plots show the degradome signatures within the local regions surrounding the miRNA binding sites (marked by blue bars). Those prominent signatures mapped to the 9th to 12th nucleotides of the regulatory miRNAs were regarded as the cleavage signals. The miRNA—target al.ignments were performed by TargetFinder [36]. The penalties measure the binding affinity of the miRNA—target pairs in different *Arabidopsis* accessions or rice varieties (a higher penalty indicates the lower binding affinity). The SNPs were highlighted in red

MBS prediction for all of the miRBase-registered miRNAs (release 22), in order to obtain the miRNA—target pair lists of Col-0 and NIP respectively. Treating these pairs as the references, we searched for the polymorphic sites allocated either to the miRNAs or to the MBSs. For each pair containing the polymorphic site(s), the change of the miRNA—target affinity was assessed by TargetFinder [36]. For *Arabidopsis*, the miRNA—target pairs met the following criterion would be retained: the difference of the penalty between Col-0 and the other accession was greater than or equal to "2". Compared to Col-0, the increased penalty of a pair indicates the reduced miRNA—target affinity, while the decreased penalty indicates the elevated affinity. In *Arabidopsis*, a total of 205, 113, 240, 266, 169 and 244 miRNA—target pairs with significant changes of interaction affinity were identified in Altai-5, ICE73, Kas-2, Kondara, Se-0 and Zal-1, respectively (Table S11). The same criterion was applied for rice, by using NIP as the reference. As a result, 909, 1,105, 1,110, 1,232 and 101 pairs with significant changes of binding affinity were identified in 93 – 11, DY5, MBWX, NTH and SW295, respectively (Table S11). All of these miRNA—target pairs could be classified into three classes according to the locations of the polymorphic sites: (1) the pairs with the polymorphic sites allocated to the mature miRNAs; (2) the pairs with the polymorphic sites allocated to the MBSs; and (3) the pairs with the polymorphic sites allocated to both regions. Interestingly, except for SW295, the numbers of the miRNA—target pairs belonging to the second class are much higher than those of the other two classes in both *Arabidopsis* and rice.

Then, the miRNA—target pairs identified above were subject to degradome-based validation as described previously [5, 6]. For each validated pair, the site-specific cleavage signal intensity was compared between the polymorphic site-containing accessions and Col-0 in *Arabidopsis*, and between the polymorphic site-containing varieties and NIP in rice, respectively. Only those pairs with cleavage signals well correlated with the changes of the calculated penalty-based interaction affinity were retained. That is, the reduced affinity of a pair in a subspecies will theoretically lead to the weakened cleavage signals in this subspecies, while the increased affinity will enhance the cleavages accordingly. As a result, a total of 14 and 12 miRNA—target pairs were discovered in *Arabidopsis* and rice respectively (Fig. 6 and Fig. S4). For example, an MBS of ath-miR397b was discovered on the transcript AT2G29130.1 in Col-0, and the binding affinity of this pair was measured with the TargetFinder-calculated penalty "3.5". Degradome-based cleavage signal could be hardly detected in Col-0, indicating the weak or fake interaction between ath-miR397b and AT2G29130.1. However, a conserved SNP "U to G" was discovered at the 13th nucleotide of ath-miR397b in Altai-5, Kondara and Zal-1, respectively. This SNP leads to a significant elevation of the affinity between ath-miR397b and AT2G29130.1, as indicated by the penalty "1.5". Accordingly, the degradome-based cleavage signals could be detected in all of the three accessions (Fig. 6A). Based on the annotation, *AT2G29130* was responsive to water deprivation. To seek for the experimental evidence, the gene expression data available at the eFP Browser [35] were used. Expectedly, the expression of *AT2G29130* was significantly induced under the drought treatment (Fig. S5A). Notably, the annual precipitation in the natural habitats of Altai-5, Kondara and Zal-1 is 170 mm, 558 mm and 302 mm respectively, while that of Col-0 is much higher (1,012 mm). Thus, the SNP-induced differential regulation of *AT2G29130* may contribute to the distinct water tolerance in different accessions. In another case, an MBS of ath-miR419 was identified on the transcript AT3G13690.1. However, the miRNA—target interaction affinity is relatively low in Col-0, according to the calculated penalty "5". In Altai-5, a "C to A" SNP site on ath-miR419 results in the increased affinity between the miRNA and AT3G13690.1. Based on the degradome-seq data, the cleavage signal intensity is much higher in Altai-5 than that in Col-0 (Fig. 6B). The eFP expression data show that the expression of *AT3G13690* is repressed under multiple stress treatments, such as cold, heat, salt and drought (Fig. S5B). In this regard, *AT3G13690* plays a potential role in stress signal transduction. Also in *Arabidopsis*, an MBS of ath-miR414 was detected on the transcript AT5G15140.1. In Col-0, the ath-miR414—AT5G15140.1 interaction was relatively weak based on

the calculated penalty "5.5". However in Zal-1, two SNPs discovered on the MBS and the mature miRNA greatly enhanced the miRNA—target interaction as reflected by the penalty "3". Accordingly, the stronger cleavage signals were detected from the degradome-seq library of Zal-1 (Fig. 6C). It was noticed that the annual precipitation in the natural habitat of Zal-1 (302 mm) is much less than Col-0 (1,012 mm), and the annual mean temperature of Zal-1 (-1.1℃) is much lower than Col-0 (13.0℃). Intriguingly, based on the expression data available at the eFP Browser, the expression of *AT5G15140* was significantly repressed under the drought and heat treatments respectively (Fig. S5C), pointing to its biological role in water and temperature sensing.

In rice, an MBS of osa-miR1846c-5p was detected on LOC_Os03g10620.1 in NIP, and the miRNA—target interaction is supported by the degradome-seq data. However in 93 – 11 and DY5, the SNPs discovered on the MBS and the mature miRNA significantly weaken the interaction, since the penalty is increased from "5" to "7". No cleavage signal was detected in the two varieties accordingly (Fig. 6D). The expression data retrieved from RGAP7.0 [26] show that *LOC_Os03g10620* can be repressed under the salt and drought stress treatments respectively (Fig. S5D). Also in rice, the interaction between osa-miR159f and LOC_Os01g59660.1 was predicted in both NIP and 93 – 11. Compared to NIP, the miRNA—target interaction in 93 – 11 is interfered by a SNP site on osa-miR159f. Based on the degradome-seq data, the cleavage signal intensity is reduced in 93 – 11 accordingly (Fig. 6E). Functional annotations indicate that *LOC_Os01g59660* is potentially involved in stimulus response and reproduction. Consistently, the expression data retrieved from RGAP7.0 show that the gene expression can be greatly inhibited under the cold and heat treatments respectively, and is highly active in inflorescences (Fig. S5E).

For another case in rice, the interaction between osa-miR2880 and LOC_Os03g40920.1 was detected in NIP, which was supported by the degradome-based cleavage signal. However, in DY5, three SNPs on the MBS lead to the decreased interaction affinity of this pair (Fig. 6F). As reflected by the expression data retrieved from RGAP7.0, *LOC_Os03g40920* is responsive to multiple stress treatments, such as heat, salt and drought (Fig. S5F).

## Discussion

In the present study, 24–35 differentially expressed miRNA precursors and 36–110 differentially expressed mature miRNAs were identified in the six *Arabidopsis* accessions, as compared to Col-0 (Fig. 1A and B). And, 19–40 differentially expressed precursors and 8–101 mature miRNAs were identified in the five rice varieties, as compared to Nipponbare (Fig. 1C and D). Considering the large discrepancy of the natural habitat parameters among distinct subspecies, the differential miRNA expression may have a potential contribution to the different environmental adaptability of the subspecies investigated. On the other hand, our analysis shows that in most cases, the expression pattern of a mature miRNA correlates well with that of its precursor (Fig. 1E and F), supporting the notion that the mature miRNA production is highly dependent on the expression level of the precursor.

Based on the MeRIP-seq data, several miRNA precursors were discovered to be differentially methylated in distinct subspecies of *Arabidopsis* and rice (Fig. 2E and F). Among these cases, some differential $m^6A$ peaks are overlapped with the miRBase-registered precursors, while some peaks are overlapped with the 200-nt extensions of the precursors. Similar to the protein-coding genes, the plant miRNA genes are transcribed by RNA polymerase II [3]. The poly(A)-tailed primary transcripts termed pri-miRNAs are unstable and hard to be cloned. Instead, the secondary transcripts with the stem-loop structures, named as pre-miRNAs, are usually registered in miRBase [32]. Thus, our observation indicates that differential $m^6A$ modification may occur on the primary transcripts of the miRNA genes. On the other hand, it is well-known that $m^6A$ modification plays an important regulatory role in transcript stability [34] and miRNA processing [11–13] in plants. Interestingly, our results show that both hyper- and hypo-methylation will lead to the up- or down-regulation of miRNA expression (Fig. 3).

The SNPs were identified to have a significant impact on the stem-loop structures of the miRNA precursors. Supported by the sequencing data, these SNP sites were suggested to have a direct impact on the precursor stability and the mature miRNA levels. Interestingly, in several cases, the SNP sites along with the miRNA expression patterns are highly conserved among the subspecies investigated (Fig. 4). Moreover, the SNP sites discovered on the MBSs or the mature miRNAs, have an obvious influence on the interaction affinity of the miRNA—target pairs (Fig. 6 and Fig. S4). Intriguingly, based on the tissue- and treatment-specific expression data, many of these target genes were indicted to have important biological functions related to organ development, reproduction or stress responses. Indeed, some of these SNP sites were observed to be evolutionarily conserved among several *Arabidopsis* accessions or rice varieties. Degradome sequencing data analysis shows that the high abundance of a mature miRNA likely leads to intense target cleavages in a specific subspecies. Interestingly, based on the tissue-specific expression data, most of the target genes are likely to be involved in reproduction.

## Conclusions

Summarily, based on the study on distinct *Arabidopsis* accessions and rice varieties, our results show the great influence of the genomic and epitranscriptomic variations on the expression and regulatory activity of the miRNA genes. The data provide the molecular hints potentially linked to the different developmental processes and environmental responses among the plant subspecies with distinct natural habitats.

### Abbreviations

| | |
|---|---|
| miRNA | microRNA |
| SNP | single nucleotide polymorphism |
| $m^6A$ | $N^6$-methyladenosine |
| AGO1 | Argonaute 1 |
| MBS | miRNA binding site |
| RACE | rapid amplification of cDNA ends |
| degradome-seq | degradome sequencing |
| sRNA-seq | small RNA sequencing |
| RPM | reads per million |
| RIN | RNA integrity number |
| TAIR | The Arabidopsis Information Resource |
| RGAP | The rice genome annotation project |
| IGV | Integrative genomics viewer |
| FPKM | fragments per kilobase of exon per million fragments mapped |
| SRA | Sequence Read Archive |
| eFP | electronic Fluorescent Pictograph |

## Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12870-026-08516-8.

Supplementary Material 1.

### Acknowledgements

We thank the National Mid-term Genebank for Rice of China National Rice Research Institute for providing the rice germplasm collection (Nipponbare, 93-11, DANG YU 5 HAO, MIN BEI WAN XIAN, NAN TE HAO, Suweon 295). We thank Hidetaka Ito (Faculty of Science, Hokkaido University) for providing the different accessions of Arabidopsis (Col-0, Altai-5, ICE73, Kas-2, Kondara, Se-0 and Zal-1).

### Authors' contributions

Conception and methods, original manuscript preparation: Y.M., X.M. and C.S.; data analyses and testing of data accuracy: Z.F., J.Y., J.W. X.W. and W.W.; prepared the plant materials: Z.J. and J.Y.; Funding acquisition: Y.M.; All authors read and approved the final version of the manuscript.

### Funding

This study was supported by grants from the National Natural Science Foundation of China [32370707].

### Data availability

All of the sequencing data reported in this study has been deposited in the Sequence Read Archive (SRA) database under the accession IDs PRJNA1229131 (sRNA-seq data of the seven Arabidopsis accessions), PRJNA1230804 (sRNA-seq data of the six rice varieties), PRJNA1229083 (degradome-seq data of the seven Arabidopsis accessions), PRJNA1231718 (degradome-seq data of the six rice varieties), PRJNA1163906 (m6A MeRIP-seq data of the seven Arabidopsis accessions), PRJNA1230748 (m6A MeRIP-seq data of the six rice varieties), PRJNA1232210 (RNA-seq data of the seven Arabidopsis accessions) and PRJNA1231720 (RNA-seq data of the six rice varieties). The sequence polymorphism datasets of the Arabidopsis accessions were reported by [1] and available at http://1001genomes.org. The sequence polymorphism datasets of the rice varieties were reported by [2] and available at https://github.com/zhuochenbioinfo/3KRG-HAP. The tissue- and treatmen t-specific gene expression data of Arabidopsis and rice are publicly available at the eFP Browser (https://www.bar.utoronto.ca/) [35] and RGAP7.0 (https://rice.uga.edu/) [26], respectively. The geographic distribution information (longitude and latitude) of the seven Arabidopsis ecotypes was retrieved from https://1001genomes.org/accessions.html [1], and the information of the six rice varieties was available in a previous report [37]. The latitude and longitude coordinates were used to query the WorldClim 2.1 database (http://www.worldclim.org) [38] for the environmental parameters of the plant natural habitats.

## Declarations

### Ethics approval and consent to participate

Field and laboratory studies were conducted by local legislation. This article does not contain any studies with human participants or animals and does not involve any endangered or protected species. The plant materials sampled and experiments performed in this research complied with institutional, national, and international guidelines and legislation.

### Consent for publication

Not applicable.

### Competing interests

The authors declare no competing interests.

Received: 20 October 2025 / Accepted: 3 March 2026
Published online: 06 March 2026

### References

1. Consortium TG. 1,135 Genomes Reveal the Global Pattern of Polymorphism in Arabidopsis thaliana. Cell. 2016;166(2):481–91.
2. Wang W, Mauleon R, Hu Z, Chebotarov D, Tai S, Wu Z, Li M, Zheng T, Fuentes RR, Zhang F, et al. Genomic variation in 3,010 diverse accessions of Asian cultivated rice. Nature. 2018;557(7703):43–9.
3. Voinnet O. Origin, biogenesis, and activity of plant microRNAs. Cell. 2009;136(4):669–87.
4. Llave C, Xie Z, Kasschau KD, Carrington JC. Cleavage of Scarecrow-like mRNA targets directed by a class of Arabidopsis miRNA. Science. 2002;297(5589):2053–6.
5. Meng Y, Shao C, Chen M. Toward microRNA-mediated gene regulatory networks in plants. Brief Bioinform. 2011;12(6):645–59.
6. Shao C, Chen M, Meng Y. A reversed framework for the identification of microRNA-target pairs in plants. Brief Bioinform. 2013;14(3):293–301.
7. Chen X. Small RNAs and their roles in plant development. Annu Rev Cell Dev Biol. 2009;25:21–44.
8. Todesco M, Balasubramanian S, Cao J, Ott F, Sureshkumar S, Schneeberger K, Meyer RC, Altmann T, Weigel D. Natural variation in biogenesis efficiency of individual Arabidopsis thaliana microRNAs. Curr Biol. 2012;22(2):166–70.
9. Rojas AML, Drusin SI, Chorostecki U, Mateos JL, Moro B, Bologna NG, Bresso EG, Schapire A, Rasia RM, Moreno DM, et al. Identification of key sequence features required for microRNA biogenesis in plants. Nat Commun. 2020;11(1):5320.
10. Alarcon CR, Lee H, Goodarzi H, Halberg N, Tavazoie SF. N6-methyladenosine marks primary microRNAs for processing. Nature. 2015;519(7544):482–5.
11. Bai H, Dai Y, Fan P, Zhou Y, Wang X, Chen J, Jiao Y, Du C, Huang Z, Xie Y, et al. The METHYLTRANSFERASE B-SERRATE interaction mediates the reciprocal regulation of microRNA biogenesis and RNA m(6)A modification. J Integr Plant Biol. 2024;66(12):2613–31.
12. Bhat SS, Bielewicz D, Gulanicz T, Bodi Z, Yu X, Anderson SJ, Szewc L, Bajczyk M, Dolata J, Grzelak N, et al. mRNA adenosine methylase (MTA) deposits $m^6A$ on pri-miRNAs to modulate miRNA biogenesis in Arabidopsis thaliana. Proc Natl Acad Sci U S A. 2020;117(35):21785–95.
13. Zhong S, Li X, Li C, Bai H, Chen J, Gan L, Zhu J, Oh T, Yan X, Zhu J, et al. SERRATE drives phase separation behaviours to regulate m6A modification and miRNA biogenesis. Nat Cell Biol. 2024;26(12):2129–43.
14. Meng Y, Gou L, Chen D, Mao C, Jin Y, Wu P, Chen M. PmiRKB: a plant microRNA knowledge base. Nucleic Acids Res. 2011;39(Database issue):D181–187.
15. Ehrenreich IM, Purugganan MD. Sequence variation of MicroRNAs and their binding sites in Arabidopsis. Plant Physiol. 2008;146(4):1974–82.

16. Wu X, Wang X, Chen W, Liu X, Lin Y, Wang F, Liu L, Meng Y. A microRNA-microRNA crosstalk network inferred from genome-wide single nucleotide polymorphism variants in natural populations of Arabidopsis thaliana. Front Plant Sci. 2022;13:958520.
17. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018;34(17):i884–90.
18. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol. 2019;37(8):907–15.
19. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26(1):139–40.
20. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics. 2016;32(18):2847–9.
21. Lyu F, Han FR, Ge CL, Mao WK, Chen L, Hu HP, et al. OmicStudio: a composable bioinformatics cloud platform with real-time feedback that can generate high-quality graphs for publication. Imeta. 2023;2(1):e85.
22. Dai XB, Zhuang ZH, Zhao PXC. psRNATarget: a plant small RNA target analysis server (2017 release). Nucleic Acids Res. 2018;46(W1):W49–54.
23. Zhang Y. miRU: an automated plant miRNA target prediction server. Nucleic Acids Res. 2005;33(Web Server issue):W701–704.
24. Hofacker IL. RNA secondary structure analysis using the Vienna RNA package. Curr Protoc Bioinformatics. 2009, Chap. 12:12 12 11–12 12 16.
25. Reiser L, Bakker E, Subramaniam S, Chen X, Sawant S, Khosa K, et al The Arabidopsis information resource in 2024. Genetics. 2024;227(1):iyae027.
26. Hamilton JP, Li CX, Buell CR. The rice genome annotation project: an updated database for mining the rice genome. Nucleic Acids Res. 2024;53(D1):D1614–22.
27. Meng J, Cui X, Rao MK, Chen Y, Huang Y. Exome-based analysis for RNA epigenome sequencing data. Bioinformatics. 2013;29(12):1565–7.
28. Meng J, Lu Z, Liu H, Zhang L, Zhang S, Chen Y, Rao MK, Huang Y. A protocol for RNA methylation differential analysis with MeRIP-Seq data and exomePeak R/Bioconductor package. Methods. 2014;69(3):274–81.
29. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. Integrative genomics viewer. Nat Biotechnol. 2011;29(1):24–6.
30. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014;30(7):923–30.
31. Song XW, Li Y, Cao XF, Qi YJ. MicroRNAs and Their Regulatory Roles in Plant-Environment Interactions. Annu Rev Plant Biol. 2019;70:489–525.
32. Kozomara A, Birgaoanu M, Griffiths-Jones S. miRBase: from microRNA sequences to function. Nucleic Acids Res. 2019;47(D1):D155–62.
33. Wang GQ, Li HX, Ye C, He KY, Liu S, Jiang BC, et al. Quantitative profiling of $m^6A$ at single base resolution across the life cycle of rice and Arabidopsis. Nat Commun. 2024;15(1):4881.
34. Yue H, Nie X, Yan Z, Weining S. N6-methyladenosine regulatory machinery in plants: composition, function and evolution. Plant Biotechnol J. 2019;17(7):1194–208.
35. Winter D, Vinegar B, Nahal H, Ammar R, Wilson GV, Provart NJ. An Electronic Fluorescent Pictograph browser for exploring and analyzing large-scale biological data sets. PLoS ONE. 2007;2(8):e718.
36. Allen E, Xie Z, Gustafson AM, Carrington JC. microRNA-directed phasing during trans-acting siRNA biogenesis in plants. Cell. 2005;121(2):207–21.
37. Zheng X, Pang H, Wang J, Yao X, Song Y, Li F, Lou D, Ge J, Zhao Z, Qiao W, et al. Genomic signatures of domestication and adaptation during geographical expansions of rice cultivation. Plant Biotechnol J. 2022;20(1):16–8.
38. Fick SE, Hijmans RJ. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. Int J Climatol. 2017;37(12):4302–15.

## Publisher's note